의학 박사 학위논문

환자 유래 대장암 세포주/오가노이드 생체은행을 활용한 대장암의 이질성 분석과 이에 따른 항암제 반응성 변화

A Living Biobank of Patient-Derived Colorectal Cancer Organoids and Cell Lines Captures Heterogeneity and Enables Preclinical Therapeutic Screening

2020년 8월

서울대학교 대학원

의과학과 의과학 전공

김 순 찬

A Living Biobank of Patient-Derived Colorectal Cancer Organoids and Cell Lines Captures Heterogeneity and Enables Preclinical Therapeutic Screening

By Soon-Chan Kim

(Directed by Ja-Lok Ku, D.V.M., Ph.D.)

A Thesis Submitted to the Department of Biomedical Sciences in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Medicine (Biomedical Sciences) at Seoul National University College of Medicine

June, 2020

Approved by thesis committee

Professor _____ Chairman

Professor _____ Vice Chairman

Professor _____

Professor _____

Professor _____

Abstract

김 순 찬 (Soon-Chan Kim)

의과학과 의과학 전공 (Biomedical Sciences Major)

The Graduate School of Medicine

Seoul National University

Intra-tumor heterogeneity stands for one of the main difficulties in the treatment of cancer. Since a single tumor mass consists of multiple sub-clones, targeting partial clones eventually causes drug resistance and leads to loco-regional metastasis. Therefore, a standard protocol and tool for assessing the intra-tumor heterogeneity are required for the prevention of metastasis and drug resistance. In this research, we determine patient derived tumor organoids (PDOs) by an effective way of a predictive means for analyzing intra-tumor heterogeneity of genomic and transcriptomic variances. Colorectal cancer (CRC) is extremely heterogeneous disease in terms of both a clinical and molecular perspective. Distinctive molecular characters, such as microsatellite instability (MSI), have been identified to contribute biologically distinct types of CRC with specific clinical courses. Recent study determined that

CRC can be categorized into four sub-types which is named consensus molecular subtypes (CMS). Each type exhibits an exclusive molecular characters and genetic expressions. Nevertheless, the gap between clinical applications and the fundamental research has not been closed due to the lack of reliable models. This study is dedicated to configure the heterogeneity of CRC in terms of molecular characters and gene expression patterns with the presence of CMSs in a panel of CRC cell lines and PDOs. Here, we designate a biobank of cancer cell lines and PDOs that recapitulate the histopathological and molecular variances of human CRC. Our platform contributes to a repository of heterogeneous CRC cell lines and PDOs. The majority of our model was allocated to a specific type of CMSs regardless of the absence of stromal components. We evaluated that our CMS classification with high throughput drug screening. Our resource furnishes researchers with a platform to study CRC with evident heterogeneity.

Keywords: CRC, Organoid, Heterogeneity, Drug, CMS

Student number: 2014-22004

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed human malignancy worldwide and represents the second most common cause of tumor-associated mortalities in Korea (1, 2). Regardless of the clinical accomplishments in early detection and prevention that brought about a general decrease of CRC risk (3), acquired resistance to certain chemotherapy remains major causes of CRC-associated morbidity, with scarcely remedial options accessible (4). Intratumor heterogeneity (ITH) and cancer clonal evolution have pulled in expanding attention since ITH acquired throughout cancer progression seemingly contributes to the increased therapeutic resistance and therefore lethal outcome of malignancy (5). Genetic intra-tumor heterogeneity has been reported in several types of solid tumors such as renal (6), breast (7), esophageal (8), lung (9, 10), ovarian (11, 12), prostate (13, 14), and pancreatic (15) tumors. Multiregion sequencing of spatially or temporally distinct tumor regions makes it possible to follow the evolutionary trajectories of cancer cells, with pervasive somatic driver mutations positioned within the whole tumor mass, but also

with mutations that are restricted to one of sub-clones. The parental clone that has gained all the pervasive mutation branches into sub-clones, which aggregate diverse genetic changes and eventually figure ITH. An accumulated sub-clonal mutations has been reported to associate with poorer therapeutic response (16).

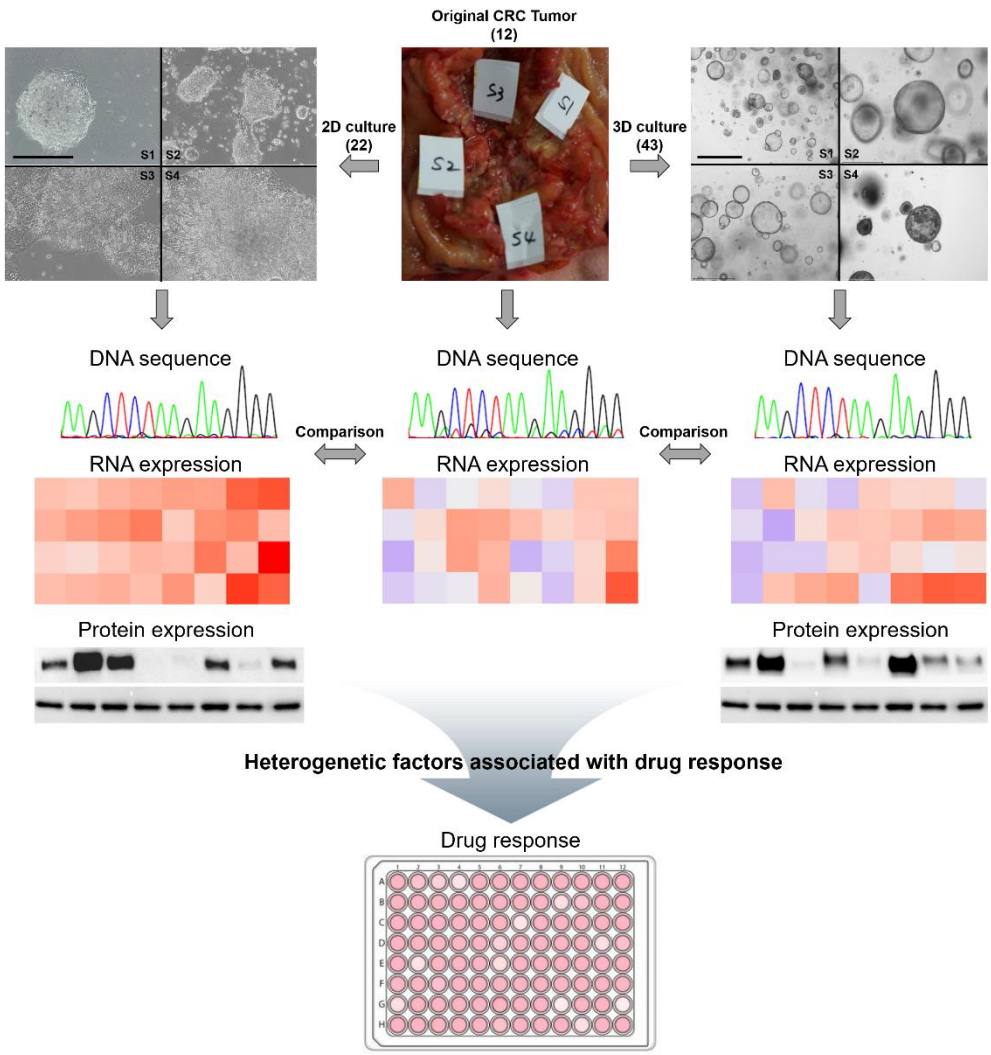
CRC represents high spatial heterogeneity and inter-patient variability in prognosis and response to certain treatments. Although some portion of these distinctions can be clarified by the serrated molecular variance (17-19) as well as microsatellite instability (MSI) (20), the multifaceted nature of CRC make it inadequate to understand the biology behind heterogeneity. For more comprehensive analysis, gene expression based molecular subtypes were defined and characterized (21-24). To determine discrepancies among the documented gene expression-based CRC subtyping, four consensus molecular subtypes (CMSs) were proposed with distinctive characteristics. CMS1 reflects microsatellite instability tumors that have immune activation with relatively good prognosis. CMS2 is canonical subtype including WNT and MYC signaling activation with epithelial features. CMS3 has

metabolic features with KRAS-mutated tumors. Finally, CMS4 is enriched for mesenchymal subtypes encompassing prominent stromal invasion and activation of transforming growth factor-β and vascular endothelial growth factor receptor pathways. It shows worst prognosis compared to other subtypes (25). CMS stratification can also be used to estimate therapeutic outcome (26). For instance, CMS type 2 CRCs have displayed better response towards oxaliplatin, whereas other CMS types were resistant. Also, irinotecan exhibited specifically better response to CMS type 4 metastatic cancers (27). Nevertheless, the gap between clinical applications and the fundamental research has not been closed due to the lack of reliable models reflecting the diverse CMSs. Although a number of reports indicated CRC organoids are capable of capturing the histological and molecular distinct of original tumor, classifying PDOs in accordance with CMSs has not been accomplished. In this perspective, the unique ability of organoid technique to recapitulate intra-tumor heterogeneity is noteworthy for comprehending tumor biology and developing personalized precision medicine. Tumor organoids progress in vitro tumor models that recapitulate the architectures of

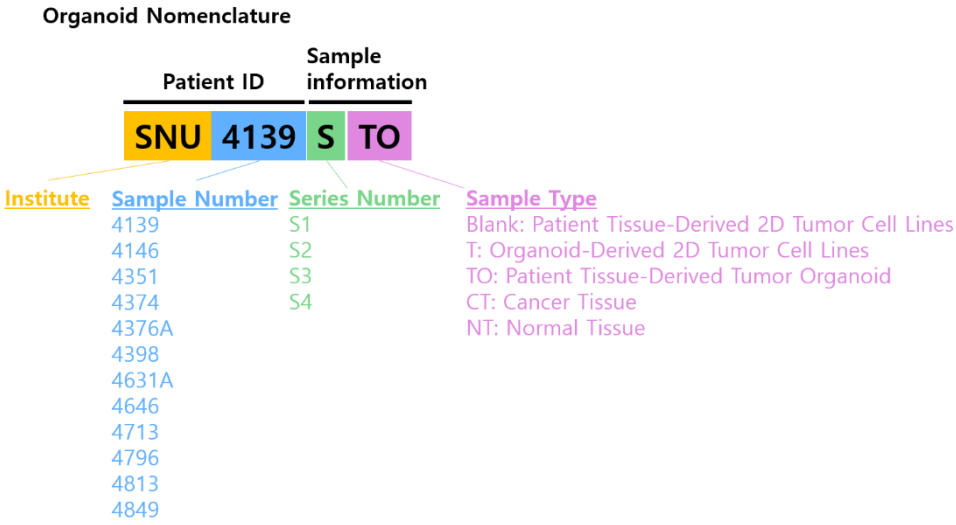
original tumorigenesis and have been recently used in the discovery of type-specific therapy and prognostic biomarkers (28, 29). Thus, the establishment and characterization of PDOs are desirable to deliver through comprehensions into molecular evolution patterns of tumors in basic research and to allow personalized anti-cancer therapy in clinical.

Material and Methods

Graphical study design and Nomenclature



continued


Organoid Nomenclature
Patient ID
Sample information
SNU
4139
S
TO
Institute
Sample Number
4139
4146
4351
4374
4376A
4398
4631A
4646
4713
4796
4813
4849
Series Number
S1
S2
S3
S4
Sample Type
Blank: Patient Tissue-Derived 2D Tumor Cell Lines
T: Organoid-Derived 2D Tumor Cell Lines
TO: Patient Tissue-Derived Tumor Organoid
CT: Cancer Tissue
NT: Normal Tissue

Ethics statement

The research protocol was reviewed and approved by the institutional review board of the Seoul National University hospitals. The study was performed in accordance with the Declaration of Helsinki. Written informed consent was obtained from all patients enrolled in this study.

Sample collection and preparation

We collected a total of 45 samples of colorectal tumors from 12 patients who underwent radical resection or endoscopic submucosal dissection at Seoul National University Hospital (Seoul, Korea). Tumor tissues were histologically diagnosed by a pathologist as carcinoma in situ. Detailed information about participants and samples is summarized in Table 1. The material information is included in last character of the sample names: "Blank", "T", "TO", "CN", and "NT" represents "Patient-Derived 2D Tumor Cell Lines", "Organoid-Derived 2D Tumor Cell Lines", "Patient-Derived Tumor Organoid", "Cancer Tissue", and "Normal Tissue" respectively. Genomic DNA was extracted from resected tumor

tissues and paired adjacent normal mucosa with All Prep DNA/RNA Mini Kit (Qiagen, Hiden, Germany).

Crypt Isolation and Culture

Crypt isolation and culture were performed according to previously documented by Hans Clevers (28) with few modifications. Normal mucosa, was stripped of the underlying muscle layer and cut into 1 – 2 mm stripes. Wash the fragments three times with chelation solution (5.6 mM Na2HPO4, 8.0 mM KH2PO4, 96.2 mM NaCl, 1.6 mM KCl, 43.4 mM Sucrose, and 54.9 mM D-Sorbitol). Dithiotreitol (DTT) was added just before use to a final concentration of 0.5 mM. Supply EDTA (2 mM final concentration) in chelation solution and incubate for 30 min at 4℃. Shake tubes vigorously to liberate the crypts. If no crypts were visible, the chelation solution/EDTA was replaced with fresh solution and the procedure was repeated until crypts were obtained. Settle the tissue fragments for 1 – 2 min and transfer the supernatant containing crypts to a new tube. Basal culture medium (advanced Dulbecco's modified Eagle medium/F12 supplemented with penicillin/streptomycin, 10 mM HEPES and Glutamax) was

added and spin down the crypts at 650 rpm for 3 min. Wash the crypts twice with basal culture medium and resuspend in Basement Membrane Extract (BME) (Cultrex (R)PC BME RGF type 2, Amsbio) and plated at different densities. After allowing the BME to solidify, add Human Intestinal Stem Cell medium (HISC) (Basal culture medium with 50% Wnt conditioned medium, 20% R-Spondin conditioned medium, 10% Noggin conditioned medium, 1 x B27, 1.25 mM n-Acetyl Cysteine, 10 mM Nicotinamide, 50 ng/ml human EGF, 10 nM Gastrin, 500 nM A83-01, 3 uM SB202190, 10 nM Prostaglandine E2 and 100 mg/ml Primocin (Vivogen) and crypts were incubated at 37℃. To make sure that as many crypts as possible were plated to ensure heterogeneity, while not overloading the seeding capacity of the BME, different densities were plated.

Tumor Isolation and Culture

Tumor isolation and culture were performed according to previously documented by Hans Clevers (28) with few modifications. Tumors were cut into pieces and two parts were processed for immunohistochemistry and DNA isolation. The remainder was cut

into smaller pieces and incubated with Collagenase II (1,5 mg/ml), Hyaluronidase (20 ug/ml) and Ly27632 (10 uM) for 30 min at 37℃ while shaking. FCS was added and the mixture was put over a 100 uM cell strainer to remove large fragments. Spin down Cells at 1,000 rpm for 3 min. Resuspending pellet in basal culture medium (advanced Dulbecco' s modified Eagle medium/F12 supplemented with penicillin/streptomycin, 10 mM HEPES and Glutamax) and spun again at 1,000 rpm 3min. The tumor material was resuspended in BME and plated at different densities. After allowing the BME to solidify, HICS minus Wnt (Basal culture medium with 20% R-Spondin conditioned medium, 10% Noggin conditioned medium, 1 x B27, 1,25 mM n-Acetyl Cysteine, 10 mM Nicotinamide, 50 ng/ml human EGF, 10 nM Gastrin, 500 nM A83-01, 3 uM SB202190, 10 nM Prostaglandine E2 and 100 mg/ml Primocin (Vivogen) was added and the cells were incubated at 37℃.

Organoid Culture

Organoid culture medium was refreshed every two days. To passage the organoids, BME was broken up by pipetting and organoids were

collected in a tube. The organoids were centrifuged at 1,000 rpm for 3 min and the medium are removed. 5 ml Triple Express (Invitrogen) was added and the organoids were incubated at 37° C for approximately 5 min. Every minute, a visual check was done to verify the size or the organoids. Care was taken not to treat the organoids to long with Triple Express. FCS and medium were added and cells were spun down at 1,500 rpm for 3 min. The pellet was taken up in BME and cells were plated in droplets of 5–10 mL each. After allowing the BME to solidify, HICS (for normal organoids) or HICS minus Wnt (for tumoroids), both supplemented with 10 uM LY27632, was added to the plates.

3D organoids Seeding/Treatment Procedure

All drug screens were performed two times. PDOs were mechanically and enzymatically dissociated into single cells by incubating in TrypLE (Gibco) for 5 to 10 min. Suspension (5 μl/well) was dispensed in clear-bottomed, white-walled 96-well plates (#3903, Corning) using an automated repeat pipet and overlaid with 200 μl of a 1:1 mixture of HISC medium and RGF basement membrane

matrix (Gibco, A14132-02). Plates are incubated at 37° C with 5% CO2 for 15 minutes to solidify the gel before addition of 20 μl of pre-warmed HISC medium to each well using an EpMotion (Eppendorf). 96 hours after seeding, 20 μl of drug containing solution is added to each well. For the control well, the mixture of HISC medium and drug-solvent solution is added.

Western Blot Analysis

7.5 × 105 cells were simultaneously seeded on a T75 flask with 15 ml of RPMI1640 media with 10% FBS and 1.1% penicillin. Cells were harvested with a cell scraper after washing with cold PBS. Whole protein was extracted with EzRIPA buffer (ATTO Co., Tokyo, JAPAN) supplied with 1% protease inhibitor and 1% phosphatase inhibitor. The volume of lysis buffer was adjusted to the number of cells collected in each vial. The protein concentration was determined by SMART™ micro BCA protein assay kit (Intron biotechnology, Gyeonggi, Korea). Equal amounts of protein were loaded on 4–15% Mini-PROTEAN TGX™ Precast Gels (BIO-RAD, Hercules, CA, USA) and blotted at 50 volts for 2 h. Proteins were then transferred to

Trans-Blot Turbo™ Transfer Pack (BIO-RAD) using Trans-Blot Turbo™ Transfer System V1.02 machine (BIO-RAD) at 2.5 Amp and 25 Volt. The membrane was incubated in 2.5% skim milk containing 0.5% Tween 20 for an hour at room temperature. Primary antibodies against EGFR (abcam, Cambridge, United Kingdom) (1:2000), pEGFR-Try1068 (Cell Signaling Technology, MA, USA) (1:1000), HER2 (abcam) (1:1000), MLH1 (Santa Cruz Biotechnology, TX, USA) (1:500), MSH2 (Santa Cruz Biotechnology) (1:500), EpCAM (Santa Cruz Biotechnology) (1:1000), E-cadherin (abcam) (1:1000), CD133 (abcam) (1:1000), ERK1/2 (Santa Cruz Biotechnology) (1:500), pERK1/2-Thr202/Tyr204 (Cell Signaling Technology) (1:1000), panAKT (Cell Signaling Technology) (1:1000), pAKT-Thr308 (Cell Signaling Technology) (1:1000), MEK1/2 (Cell Signaling Technology) (1:1000), pMEK1/2-Ser221 (Cell Signaling Technology) (1:1000), mTOR (Cell Signaling Technology) (1:1000), pmTOR-Ser2448 (Cell Signaling Technology) (1:1000), and β-actin (Cell Signaling Technology) (1:1000).

Immunocytochemistry

Four thousand cells were seeded on chambered coverglass (Thermo Fisher Scientific, Waltham, MA). The chambered coverglass was designed to be hydrophilic, and no ECM component was treated before seeding. Once 70% confluency had been reached, cells were washed with cold DPBS three times. Then, cells were fixed and permeabilized with BD Cytofix/Cytoperm (BD Science, San Jose, CA). After cells were washed with washing solution (BD Science), DPBS containing 2% FBS (GE Healthcare Life Sciences, Buckinghamshire, UK) was applied for an hour. After cells were washed with cold DPBS, CD133 antibody (Abcam, Cambridge, United Kingdom) (1:400) diluted in 0.05% of PBST was applied for 1.5 hours in room temperature. Thereafter, cells were washed with 0.05% of PBST, and Alexa 488 and Alexa 568 secondary antibodies (Thermo Fisher Scientific, Waltham, MA) (1:500) diluted in 0.05% of PBS.T were applied for an hour in room temperature. DAPI and Rhodamine-conjugated Phalloidin (Sigma-Aldrich, MO, USA) were diluted in distilled water and applied for 30 minutes in room temperature. Cells were washed with DPBS three times and applied under confocal microscope. LSM800 Confocal Laser Scanning Microscope and ZEN

software (Carl Zeiss, Oberkochen, Germany) were used to analyze images. Digital resolution, scan speed, and the number of pictures averaged were set to 1024 × 1024, 40 seconds per one channel, and 8 pictures, respectively. Diverse magnifications were used in accordance with growth patterns and sizes of cells. The intensity of each channel was fixed for comparing target protein expression between samples.

ATP assay

After 72 hours of drug treatment, 10 μl of Celltiter-Glo 3D Reagent (Promega #G968B) is added to each well followed by 5 minute of vigorous shaking. After 30 minutes incubation at room temperature and an additional minute of shaking, luminescence is measured with a Luminoskan Ascent (Thermo Scientific) over 1000 ms of integration time. Data is normalized to vehicle and plotted and EC50 values are calculated with Prism 7. For the high-throughput drug screening, DMSO is used as control. Values are normalized to vehicle.

Construction of Evolutionary Trees

Evolutionary trajectory of twelve CRC cases was traced by Treeomics algorithm (30) using whole exome sequencing data for multiregion samples. Treeomics setting was identical for all samples (sequencing error rate = 0.005, prior absent probability = 0.5, max absent VAF = 0.05, LOH frequency = 0, false discovery rate = 0.05, false-positive rate = 0.005, and absent classification minimum coverage: 100). Input parameters include read depths of both mutant and coverage genes, gene symbols, chromosomal coordinate, and substitutional patterns. Among various mutations, 298 pan-cancer driver genes (31) were selected to construct the evolutionary tree. Treeomics algorithm also provided likely driver gene mutations and built-in Cancer Gene Census List. Although MUC6 gene was included in 298 pan-cancer driver genes, the number of mutations harboring MUC6 gene outnumbered other genes. Consequently, the structure of the evolutionary tree was largely affected by the MUC6 variations. To eliminate potential bias, we excluded mutations of MUC6 genes from the input data. Sequencing artifacts were automatically adjusted by Treeomics default setting to confirm the topologic configuration of the evolutionary tree was compatible with the mutational patterns.

Sub-clonal analysis was conducted by adding "-u" parameter to input commands.

Analysis of CNVs

For the detection of Copy Number Variations (CNVs) and loss of heterozygosity (LOH) from exome sequencing data, we employed ExomeCNV package in R program (32). The final log ratio of depth of coverage was determined by the number of bases targeted by exome sequencing (targeted base) and the number of bases actually sequenced (mapped). CNV calls were expressed as 1, 2, and 3 which indicated deletion, normal and amplification respectively.

Whole-exome sequencing

Whole-exome capture was performed on all samples with the SureSelect Human All Exon V5 Kit (Agilent Technologies, Tokyo, Japan). The captured targets were subjected to sequencing using HiSeq 2500 (Illumina, San Diego, CA, USA) with the pair-end 100 bp read option for organoid samples and 200 bp read option for tissue materials. Information on read depth is provided in Supplementary

Data 2. The sequence data were processed through an in-house pipeline. Briefly, paired-end sequences are firstly mapped to the human genome, where the reference sequence is UCSC assembly hg19 (original GRCh37 from NCBI, Feb. 2009) using the mapping program BWA (version 0.7.12), and generated a mapping result file in BAM format using BWA-MEM. Then, Picard-tools (ver.1.130) were applied in order to remove PCR duplicates. The local realignment process is performed to locally realign reads with BAM files reducing those reads identically match to a position at start into a single one, using MarkDuplicates.jar, which requires reads to be sorted. By using Genome Analysis Toolkit, base quality score recalibration (BQSR) and local realignment around indels were performed. Haplotype Caller of GATK (GATKv3.4.0) was used for variant genotyping for each sample based on the BAM file previously generated (SNP and short indels candidates are detected). Somatic mutations were identified by providing the reference and sequence alignment data of tumor tissues or organoids to the MuTect2 (involved in GATK v3.8.0) with default parameters using tumor-normal mode. The matched normal tissue was not available for SNU-

4376 series, and peripheral blood mononuclear cell (PBMC) was used for somatic mutation calling. Those variants are annotated by SnpEff v4.1g, to vcf file format, filtering with dbSNP for the version of 142 and SNPs from the 1000 genome project. Then, SnpEff was applied to filter additional databases, including ESP6500, ClinVar, dbNSFP 2.9. Mutational signatures were evaluated using the Mutational Patterns R package, release 3.6.1 (33) to configure distinct footprints in genomic context for all somatic SNVs and evaluate a multitude of mutational patterns in base substitution in tumor tissues and matched cell lines/organoids.

Analysis of RNA sequencing

Paired end sequencing reads of cDNA libraries (101bp) generated from a NovaSeq6000 instrument were verified its sequence quality with FastQC v 0.11.7. For data preprocessing, low quality bases and adapter sequences in reads were trimmed using Trimmomatic v 0.38. The trimmed reads were aligned to the human genome (UCSC hg19) using HISAT v2.1.0, a splice-aware aligner. And then, transcript assembly of known transcripts, novel transcripts, and alternative

splicing transcripts was processed by StringTie v1.3.4d (34). Base on the result of that, expression abundance of transcript and gene were calculated as read count or FPKM value (Fragments Per Kilobase of exon per Million fragments mapped) per sample. Fusion detection was conducted using the default parameters for FusionCatcher v 1.05 and Defuse v0.8.16. For each program, transcriptome reads were mapped to the human genome (Ensmebl database).

Results

Establishment of Patient-Derived Colorectal Cancer Organoid Lines

We have established a living biobank of patient-derived colorectal cancer organoids that are capable of propagating in 3D culture. Surgically resected tumor tissues from twelve CRC patients were transferred directly from the operating room to the laboratory for organoid culture and DNA/RNA extraction of the original tumor. Clinicopathologic information are summarized in Table 1. The spatial sites of tumor pieces for multi-region sampling was designated (S1-S4) before preprocessing for cell line/organoid culture. Then each tumor chucks were subjected to culture for 2D cancer cell line and 3D tumor organoid. We successfully have generated 43 tumor organoids and 22 tumor cell lines corresponding to 12 different patients. Finger printing analysis indicated that each cell lines and organoids derived from a same patient shared >90% of specific loci, and not cross contaminated (Table 2). A line was considered as established when it had been successively propagated at least three times and cryopreserved. The growth rate of the organoids clone S2 from patient 4713 and clone S1 from patient 4849 diminished as

Table 1. Clinicopathological information of twelve CRC patients

SNU number	Sex/Age	MSI	c T	c N	c M	PreTx CEA	Tumor Size (cm)	Stage	Met Location	Regimen	Rec Location
SNU-4139S	F/38	MSI-L	4	1	1	16.7	8	4	liver, lung, bone, p-seeding	n/a	
SNU-4146S	M/81	MSI-L	4	1	0	0.9	5.5	3		Xeloda	
SNU-4351S	F/60	MSS	2	0	0	8.4	3.5	2		5-FU(LV)	
SNU-4374S	F/66	MSS	4	0	0	8.5	3.7	2		5-FU(LV)	
SNU-4376AS	M/39	MSI-H	4	0	1	12.2	12	4	liver	Xelox	
SNU-4398S	F/83	MSI-H	4	1	0	3.3	3.1	2		n/a	
SNU-4631AS	F/75	MSS	3	1	0	1.5	8.5	3		Xelox	
SNU-4646S	M/86	MSS	3	1	0	3.3	7	3		n/a	lung
SNU-4713S	F/86	MSS	3	1	0	3.7	5	2		n/a	
SNU-4796S	M/70	MSS	3	1	0	1.3	4.5	2		Xeloda	
SNU-4813S	M/77	MSS	3	0	0	2.3	7.5	2		n/a	
SNU-4849S	M/82	MSS	3	0	0	2.7	4	2		n/a	

Table 2. DNA fingerprinting analysis using 16STR loci for newly established 22 colorectal cancer cell lines /43 colorectal cancer organoids

Cell-Name	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	Vwa	TPOX	D18S51	Amelogenin	D5S818	FGA
SNU-4139_CancerTissueDNA	13	29,32.2	9	11	15,17	9	12	11,12	17,18	13,14.2	14	8,9	17	X	11	23
SNU-4139S1	13	29,32.2	9	11	15,17	9	12	11,12	17,18	13,14.2	14	8,9	17	X	11	23
SNU-4139S1-TO	13	29,32.2	9	11	15,17	9	12	11,12	17,18	13,14.2	14	8,9	17	X	11,12	23
SNU-4139S2-TO	13	29,32.2	9	11	15,17	9	12	11,12	17,18	13,14.2	14	8,9	17	X	11,12	23
SNU-4139S3	13	29,32.2	9	11	15,17	9	12	11,12	17,18	13,14.2	14	8,9	17	X	11	23
SNU-4139S3-TO	13	29,32.2	9	11	15,17	9	12	11,12	17,18	13,14.2	14	8,9	17	X	11	23
SNU-4139S4-TO	13	29,32.2	9	11	15,17	9	12	11,12	17,18	13,14.2	14	8,9	17	X	11	23
SNU-4146_CancerTissueDNA	11,16	29,32.2	11	11,12	16	6,9.3	8,11	12	19,24	14.2,16.2	17,18	8,11	15	X,Y	12	19,22
SNU-4146S1T	11,16	29,32.2	11	11,12	16	6,9.3	8,11	12	19,24	14.2,16.2	17,18	8,11	15	X,Y	12	19
SNU-4146S1-TO	11,16	29,32.2	11	11,12	16	6,9.3	8,11	12	19,24	14.2,16.2	17,18	8,11	15	X,Y	12	19,22
SNU-4146S2	11,16	29,32.2	11	11,12	16	6,9.3	8,11	12	19,24	14.2,16.2	17,18	8,11	15	X,Y	12	19,22
SNU-4146S2-TO	11,16	29,32.2	11	11,12	16	6,9.3	8,11	12	19,24	14.2,16.2	17,18	8,11	15	X,Y	12	19,22
SNU-4146S3	11,16	29,32.2	11	11,12	16	6,9.3	8,11	12	19,24	14.2,16.2	17,18	8,11	15	X,Y	12	22
SNU-4146S3-TO	11,16	29,32.2	11	11,12	16	6,9.3	8,11	12	19,24	14.2,16.2	17,18	8,11	15	X,Y	12	19,22
SNU-4146S4	11,16	29,32.2	11	11,12	16	6,9.3	8,11	12	19,24	14.2,16.2	17,18	8,11	15	X,Y	12	19
SNU-4146S4-TO	11,16	29,32.2	11	11,12	16	6,9.3	8,11	12	19,24	14.2,16.2	17,18	8,11	15	X,Y	12	19
SNU-4351_TilTissueDNA	11,14	30	8,10	11	15	9	8,9	9,11	17,27	13,15	17	8,10	14	X	9,11	19
SNU-4351S1-TO	11,14	30	8,10	11	15	9	8,9	9,11	17,27	13,15	17	8,10	13,14	X	9,11	19,20
SNU-4351S2-TO	11,14	30	8,10	11	15	9	8,9	9,11	17,27	13,15	17	8,10	13,14	X	9,11	19
SNU-4351S3-TO	11,14	30	8,10	11	15	9	8,9	9,11	17,27	13,15	17	8,10	13,14,15	X	9,11	19
SNU-4351S4-TO	11,14	30	8,10	11	15	9	8,9	9,11	17,27	13,15	17	8,10	13,14	X	9,11	19

Continued

Cell-Name	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	Vwa	TPOX	D18S51	Amelogenin	D5S818	FGA
SNU-4374_CancerTissueDNA	13,15	30,31.2	11	12	16,18	7,9	11,13	9,10	24,25	13,14	17,19	11	14	X	13	21,22
SNU-4374S1-TO	13,15	30,31.2	11	12	16,18	7,9	11,13	9,10	24,25	13,14	17,19	11	14	X	13	21
SNU-4376AS2-TO	13,15	30,31.2	11	12	16,18	7,9	11,13	9,10	24,24	13,14	17,19	11	14	X	13	21
SNU-4374S3-TO	13,15	30,31.2	11	12	16,18	7,9	11,13	9,10	24,25	13,14	17,19	11	14	X	13	21
SNU-4374S4-TO	13,15	30,31.2	11	12	16,18	7,9	11,13	9,10	24,25	13,14	17,19	11	14	X	13	21
SNU-4376_TilTissueDNA	12,15	30	11,12	9,12	15	9	9,10	12,13	22,23	12,13	16,17	9,11	18,19	X,Y	11.13	21,23
SNU-4376AS1T	10,14	30,31	10,11	8,12	15,16	9	9,11	11,12	22,23	12,13	17	8,11	16,22	X	12,13	18,21
SNU-4376AS1-TO	10,14	30,31	10,11	8,12	15,16	9	9,11	11,12	22,23	12,13,14	17	8,11	16,22	X	12,13	18,21
SNU-4376AS2	10,14	29,31	10,11	8,12	15,16	9	9,11	11,12	23	13	17	8,11	16,22	X	12,13	18,21
SNU-4376AS3T	10,14	30,31	10,11	8,12	15,16	9	9,11	11,12	23	12,13	17	8,11	16,22	X	12,14	18,21
SNU-4376AS3-TO	10,14	29,31	10,11	8,12	15,16	9	9,11	11,12	23	12,13	17	8,11	16,22	X	12,14	18,21
SNU-4376AS4	11,14	30	10,11	8,12	15,16	9	9,11	10,11	22,23	13	17	8,11	15,22	X,Y	12,13	18,21
SNU-4376AS4-TO	11,14	30,31	11	8,12	15,16	9	9,11	11	22,23	13	17,18	8,11	15,16,21,22	X,Y	12,13	18,21,22
SNU-4398_TilTissueDNA	12,14,13,15	30,31,31.2	12,13	11,12	16,17	7,8,9	8,11,12	9,10,11	18,19,20	12,14	17,18	8	13,14	X	12,13	21,22,23,24
SNU-4398S1	12,15	30,31.2	13	11,12	15,18	7,9	8,13	10,12	19,21	12,14	17,18	8	12	X	12,13	21,24
SNU-4398S1-TO	12,15	30,32.2	13,14	11	16,17	7,9	8,13	10,12,13	19	11,14	17,18	8	12,14	X	12,13	21,24,25
SNU-4398S2	12,15	29,31.2	12,14	11	16,17	7,9	8,12	10,11	19	12,14	17,18	8	12,13	X	12,13	21,24
SNU-4398S2-TO	12,14	30,31.2	13,14	11,12	16,18	7,9	8,12	10,11	19,20	12,14	17,19	8	13	X	12	22,25
SNU-4398S3-TO	12,15	30,31.2	12,13	11	16,17	7,9	8,13	10,12	19	12,14	17,18	8	12,14	X	12,13	20,21,25
SNU-4398S4	12,15	30,31.2	12,13	10,11	15,16	7,9	8,14,15	10,12	19,20	12,14	17,18	8	13	X	11,12	21,24
SNU-4398S4-TO	12,15	30,31.2	12,13	10,11	15,17	7,9	8,14	10,12	19,20	12,14	17,18	8	12,13	X	11,12	21,24

Continued

Cell-Name	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	Vwa	TPOX	D18S51	Amelogenin	D5S818	FGA
SNU-4631_CancerTissueDNA	13,15	29,31	10,12	10,12	15	7,9	10,11	9,12	23	12,14	17	8,11	14,19	X	11,13	26
SNU-4631AS1	13,15	29,31	10,12	10,12	15	7,9	10,11	9,12	23	12,14	17	8,11	14	X	13	26
SNU-4631AS1-TO	13,15	29,31	10,12	10,12	15	7,9	10,11	9,12	23	12,14	17	8,11	14	X	13	26
SNU-4631AS2-TO	13,15	29,31	10,12	10,12	15	9	10,11	9,12	23	12,14	17	8,11	14	X	13	26
SNU-4631AS3-TO	13,15	29,31	10,12	10,12	15	9	10,11	9,12	23	12,14	17	8,11	14	X	13	26
SNU-4631AS4	13,15	29,31	10,12	10,12	15	7,9	10,11	9,12	23	12,14	17	8,11	14	X	13	26
SNU-4631AS4-TO	13,15	29,31	10,12	10,12	15	7,9	10,11	9,12	23	12,14	17	8,11	14	X	13	26
SNU-4646_CancerTissueDNA	13,15	30,32.2	9,12	12	15,17	7,9	11	9,10	19,20	14	14	8,11	12,16	X,Y	10,11	23
SNU-4646S1T	13,15	30,32.2	9,12	12	15,17	7,9	11	9,10	20	14	14	11	16	X,Y	11	23
SNU-4646S1-TO	13,15	30,32.2	9,12	12	15,17	7,9	11	9,10	20	14	14	11	16	X,Y	11	23
SNU-4646S2T	13,15	30,32.2	9,12	12	15,17	7,9	11	9,10	20	14	14	11	16	X,Y	11	23
SNU-4646S2-TO	13,15	30,32.2	9,12	12	15,17	7,9	11	9,10	20	14	14	11	16	X,Y	11	23
SNU-4646S3T	13,15	30,32.2	9,12	12	15,17	7,9	11	9,10	20	14	14	11	16	X,Y	11	23
SNU-4646S3-TO	13,15	30,32.2	9,12	12	15,19	7,9	11	9,10	20	14	14	11	16	X,Y	11	23
SNU-4713_CancerTissueDNA	12,14	30,32.2	8,10	12,13	16,18	7,9	8,9	11,12	19,20	15.2	16,19	8,9	12,17	X,	9,10	21
SNU-4713S1	12,14	30,32.2	8,10	12,13	16,18	7,9	8	11,12	19,20	15.2	16,19	8,9	12	X	9,10	21
SNU-4713S1T	12,14	30,32.2	8,10	12,13	16,18	7,9	8,9	11,12	19,20	15.2	16,19	8,9	12	X	9,10	21
SNU-4713S1-TO	12,14	30,32.2	8,10	12,13	16,18	7,9	8,9	11,12	19,20	15.2	16,19	8,9	12	X	9,10	21
SNU-4713S2-TO	12,14	30,32.2	8,10	12,13	16,18	7,9	8,9	11,12	19	15.2	16,19	9	12	X	9,10	21
SNU-4713S3	12,14	30,32.2	8,10	12,13	16,18	7	8,9	11,12	19,20	15.2	16,19	8,9	12	X	9,10	21
SNU-4713S3-TO	12,14	30,32.2	8,10	12,13	16,18	7,9	8,9	11,12	19,20	15.2	16,19	9	12	X	9,10	21

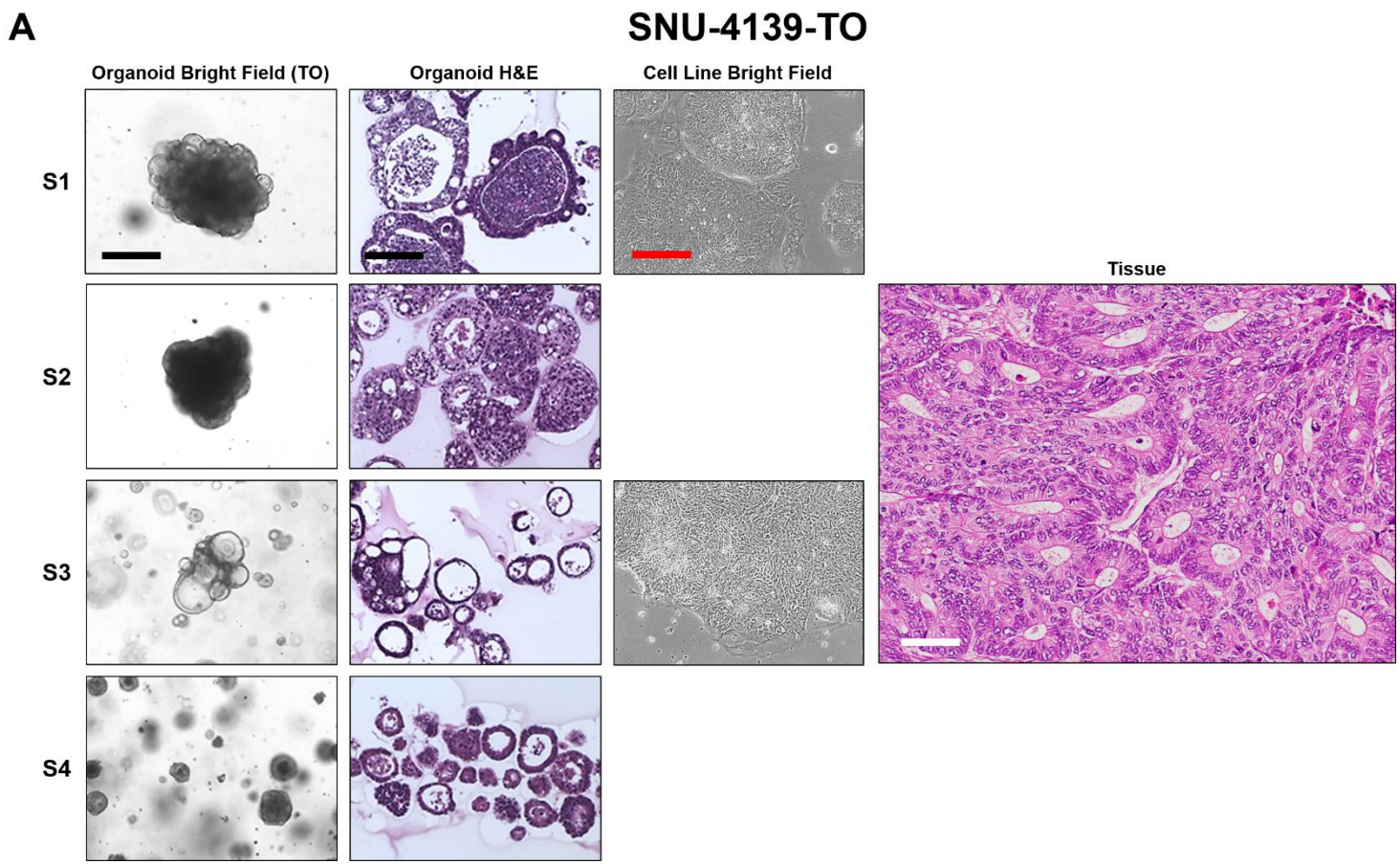
Continued

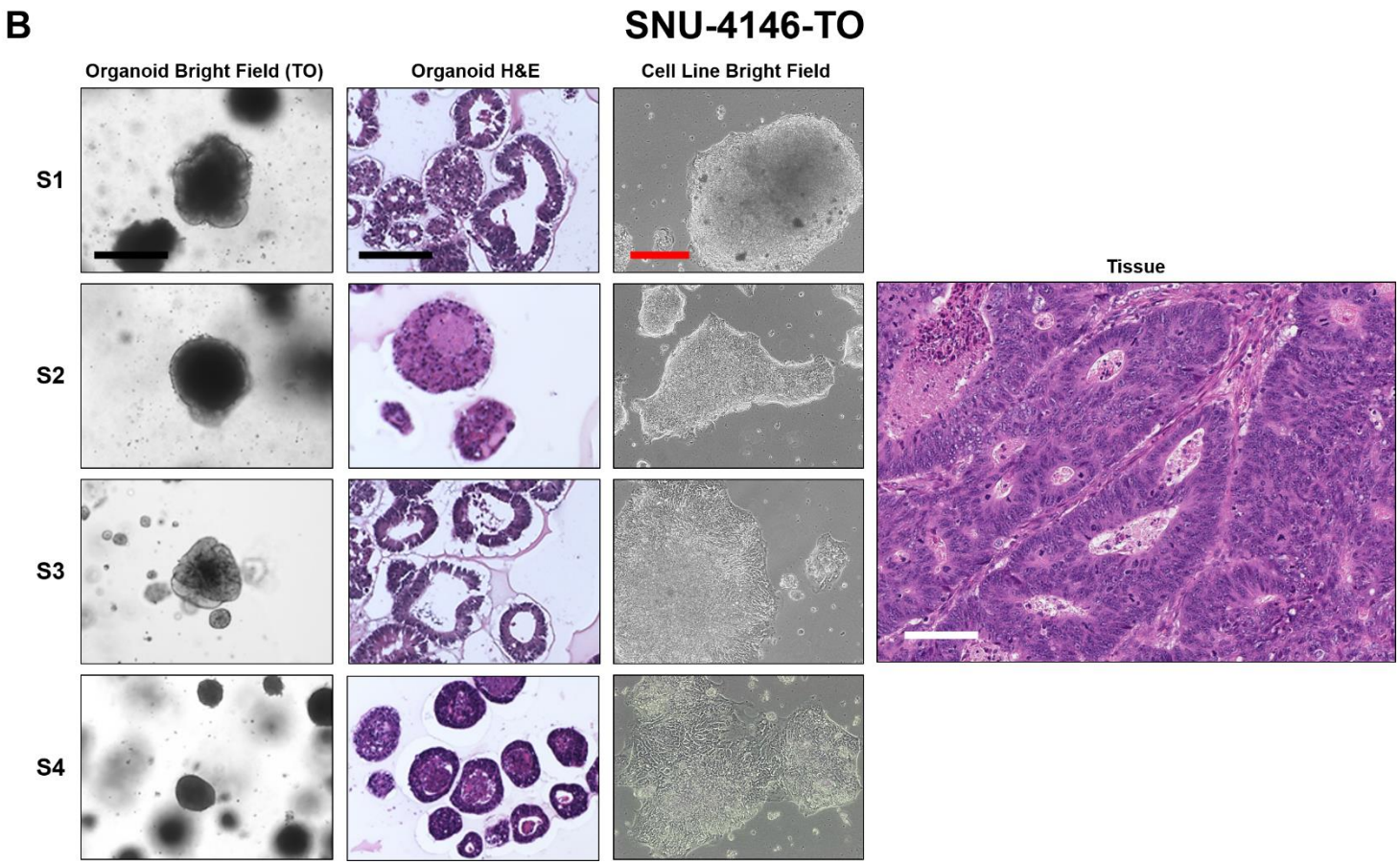
Cell-Name	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	Vwa	TPOX	D18S51	Amelogenin	D5S818	FGA
SNU-4796_CancerTissueDNA	10,15	29,32.2	11,12	12	15,17	6,9	8,11	9,11	23,24	13,13.2	14,17	8	16,19	X,Y	10,11	22,24.2
SNU-4796S1-TO	10,15	29,32.2	11,12	12	15,17	6,9	8,11	9,11	23,24	13,13.2	14,17	8	19	X,Y	11	22,24.2
SNU-4796S2	10,15	29,32.2	11,12	12	15,17	6,9	8,11	9,11	23,24	13,13.2	14,17	8	19	X,Y	11	22,24.2
SNU-4796S2-TO	10,15	29,32.2	11,12	12	15,17	6,9	8,11	9,11	23,24	13,13.2	14,17	8	19	X,Y	11	22,24.2
SNU-4796S3-TO	10,15	29,32.2	11,12	12	15,17	6,9	8,11	9	23,24	13,13.2	14,17	8	19	X,Y	11	22,24.2
SNU-4796S4-TO	10,15	29,32.2	11,12	12	15,17	6,9	8,11	9,11	23,24	13,13.2	14,17	8	19	X,Y	11	22,24.2
SNU-4813_CancerTissueDNA	14,16	29,32.2	8,10	10	15,16	9,11	8,10	9,11	18,20	13,14,14.2	14,17	11	14,19	X,Y	11,12	22.2
SNU-4813S1-TO	14,16	29,32.2	8,10	10	15,16	9,11	8,10	9,11	18,20	13,14,14.2	14,17	11	14,15	X,Y	11,12	22.2
SNU-4813S2-TO	14,16	29,32.2	8,10	10	15,16	9,11	8,10	9,11	18,20	13,14.2	14,17	11	14	X,Y	12	22.2
SNU-4813S3-TO	14,16	29,32.2	8,10	10	15,16	9,11	8,10	9,11	18,20	13,14.2	14,17	11	14	X,Y	12	22.2
SNU-4849_CancerTissueDNA	11,13	32.2	11,12	11	14,15	8,9	8,11	9,10	17,20	14,15.2	14,16	8	14,15	X,Y	11	21,24
SNU-4849S1-TO	11,13	32.2	11,12	11	15	8,9	8,11	9,10	17	14,15.2	14,16	8	15	X,Y	11	21,24
SNU-4849S2-TO	11,13	32.2	11,12	11	15	8,9	8,11	9,10	17	14,15.2	14,16	8	15	X,Y	11	21,24
SNU-4849S3-TO	11,13	32.2	11,12	11	15	8,9	8,11	9,10	17	14,15.2	14,16	8	15	X,Y	11	21,24

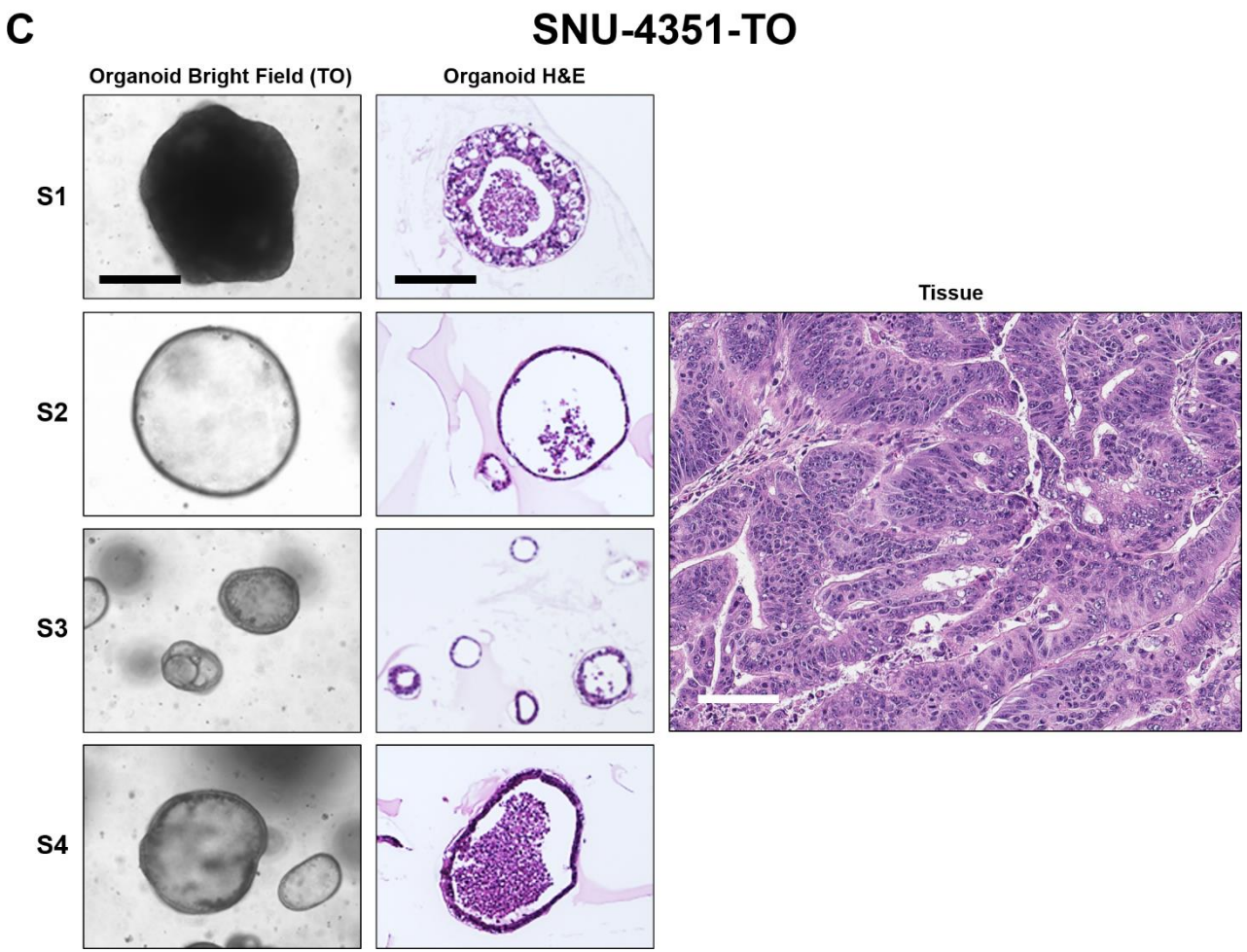
passaged, which resulted in exclusion in the HTS drug screen. The heterogeneous conformation of tumor mass was likely reflected to the variations in derivation in which areas of stromal cells or necrosis were amalgamated with viable regions. Established organoids have been consecutively passaged for up to 25 passages and have been readily cryopreserved and recovered with cell survival rate > 70%. In line with previous data (28), CRCOs varied in growth rates and morphologies (Figure 1A-L). In vitro cultivation, the organoid lines exhibited spheroidal, asymmetric and loose aggregates morphologies. The matched tumor cell lines grew as monolayers of substrate-adherent cells displaying mostly polygonal and spindle morphology. Few cell lines formed floating and adherent aggregates. The majority of tumor cells displayed a polygonal shape and had exhibited round-to-oval nuclei with prominent single-to-double nucleoli. Hematoxylin-eosin (H&E) staining of formalin fixed paraffin embedded (FFPE) organoid sections outlined that patient-derived organoids displayed sub-clonal heterogeneous morphologies ranging from thin-walled cystic structures to solid/compact structures devoid of a lumen (Figure 1A-L). Also, paraffin sections from the

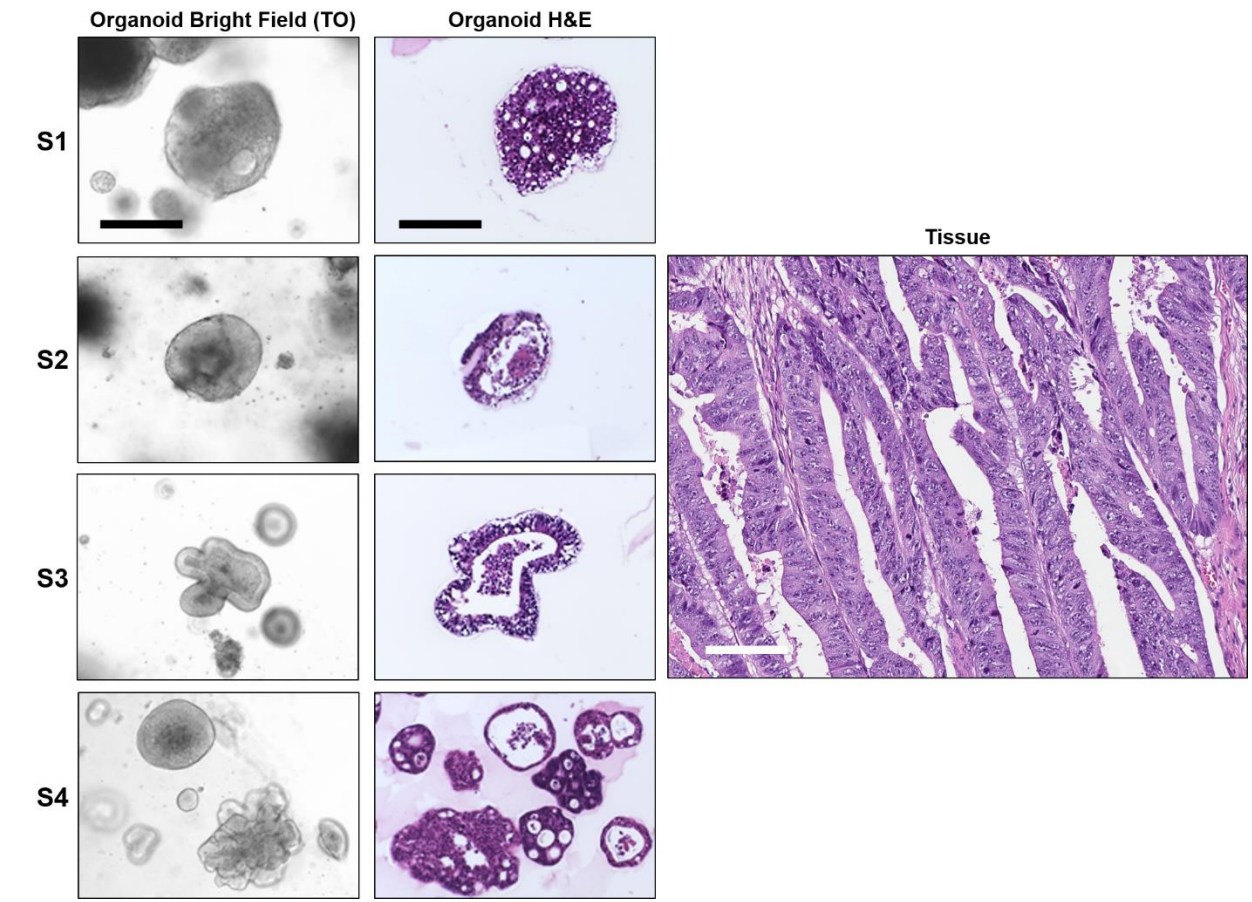
organoids as well as their corresponding parental tumors indicated strong concordance in their histopathological features (Figure 1A-L). The intra- or inter-morphological variances of established organoids were further inspected with immunocytochemistry. Figure 2A indicated that SNU-4146S1 maintained the crypt-like structure of original colorectal tumor tissue. Figure 2B displayed that there existed intra-morphological heterogeneity in SNU-4374S4-TO. One sub-clone grew as spheroidal and asymmetric shape, whereas another sub-clone retained thin-walled cystic structures with a lumen, which was already confirmed in H&E staining (Figure 1D-S4). Figure 2C showed that the internal configuration of solid/compact organoid with a thin lumen consists of not only dirty necrosis clumps but also living cells with distinct actin structure. Prominent CD133 expressing cells were detected in inner luminal region of SNU-4351S3-TO, which may suggest that the organoids maintained the cancer stem cells (Figure 2A-D). Cytokeratin 20 (CK20) and caudal type homeobox 2 (CDX2), as well as nuclear β-catenin and KI-67 were quantified and compared between matched organoids and patient tumors. Organoids retained similar presence

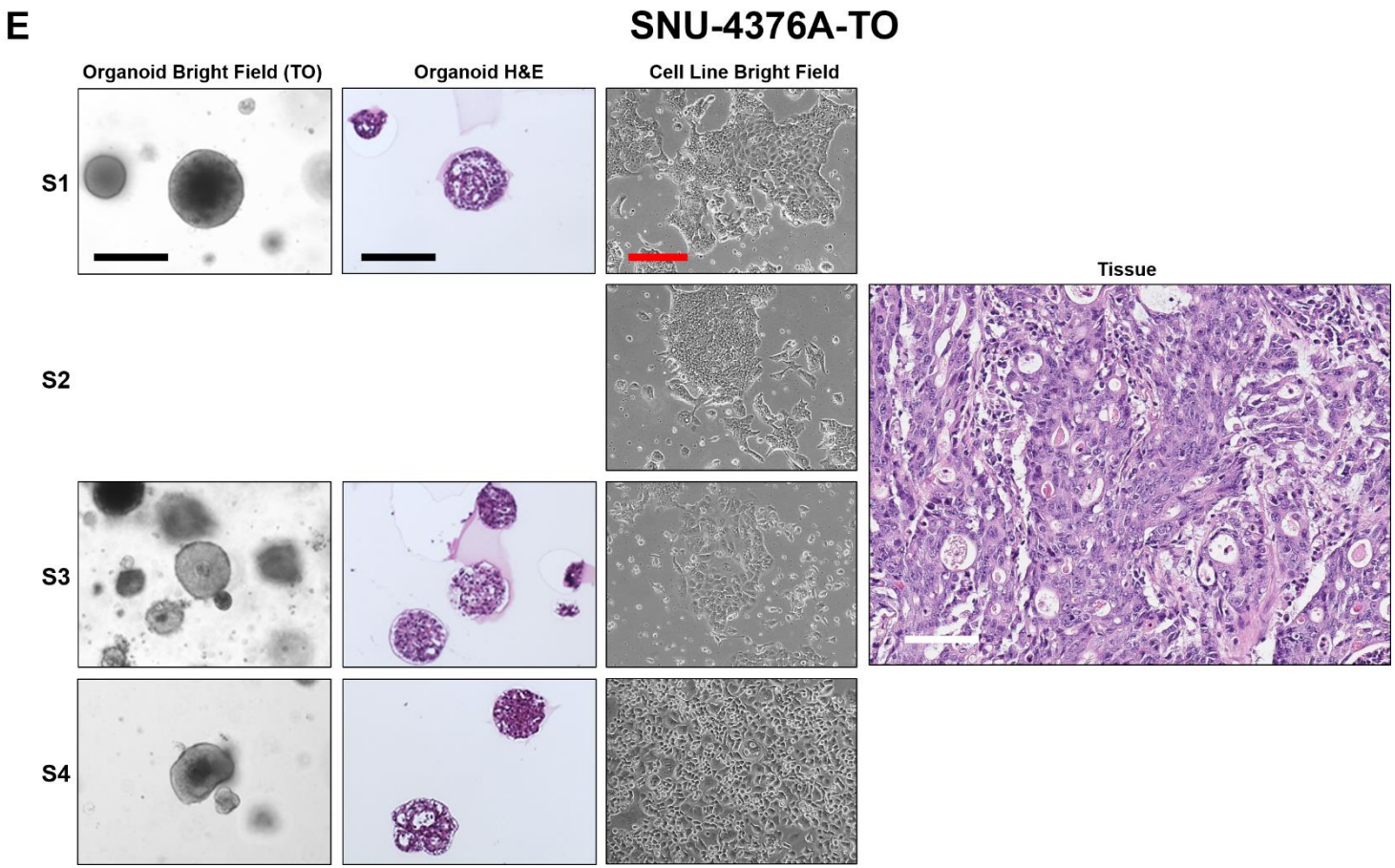
and intensity of these markers. (Figure 3A-L). MutL homolog 1 (MLH1), MutS Homolog 2 (MSH2), MutS Homolog 6 (MSH6), PMS1 Homolog 2 (PMS2) were quantified and compared between matched organoids and patient tumors. Organoids retained similar presence and intensity of these markers. SNU-4376A and SNU-4398 are deficient in MSH1 and PMS2. (Figure 4A-L)

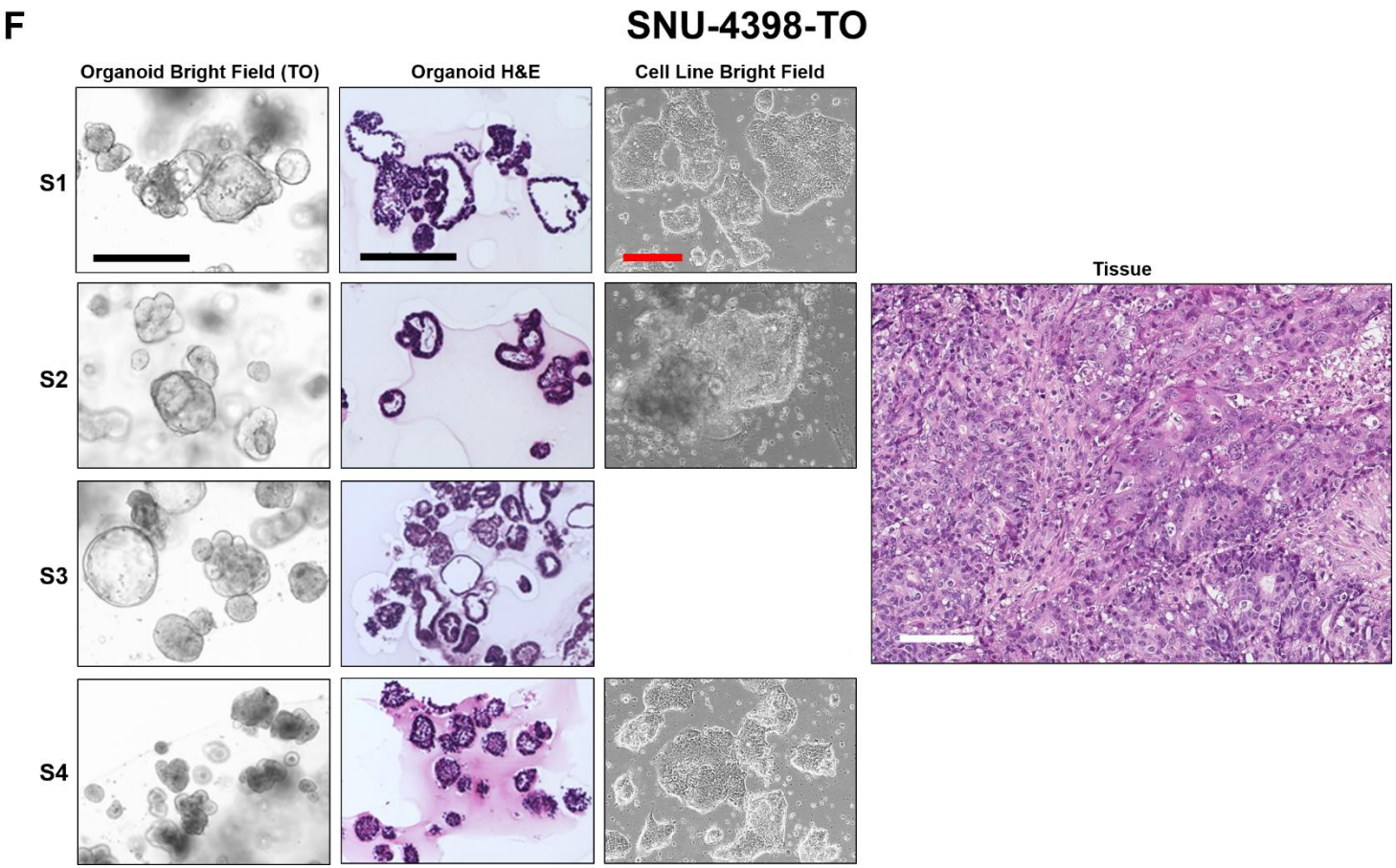

A
SNU-4139-TO
Organoid Bright Field (TO)
Organoid H&E
Cell Line Bright Field
Tissue
S1
S2
S3
S4

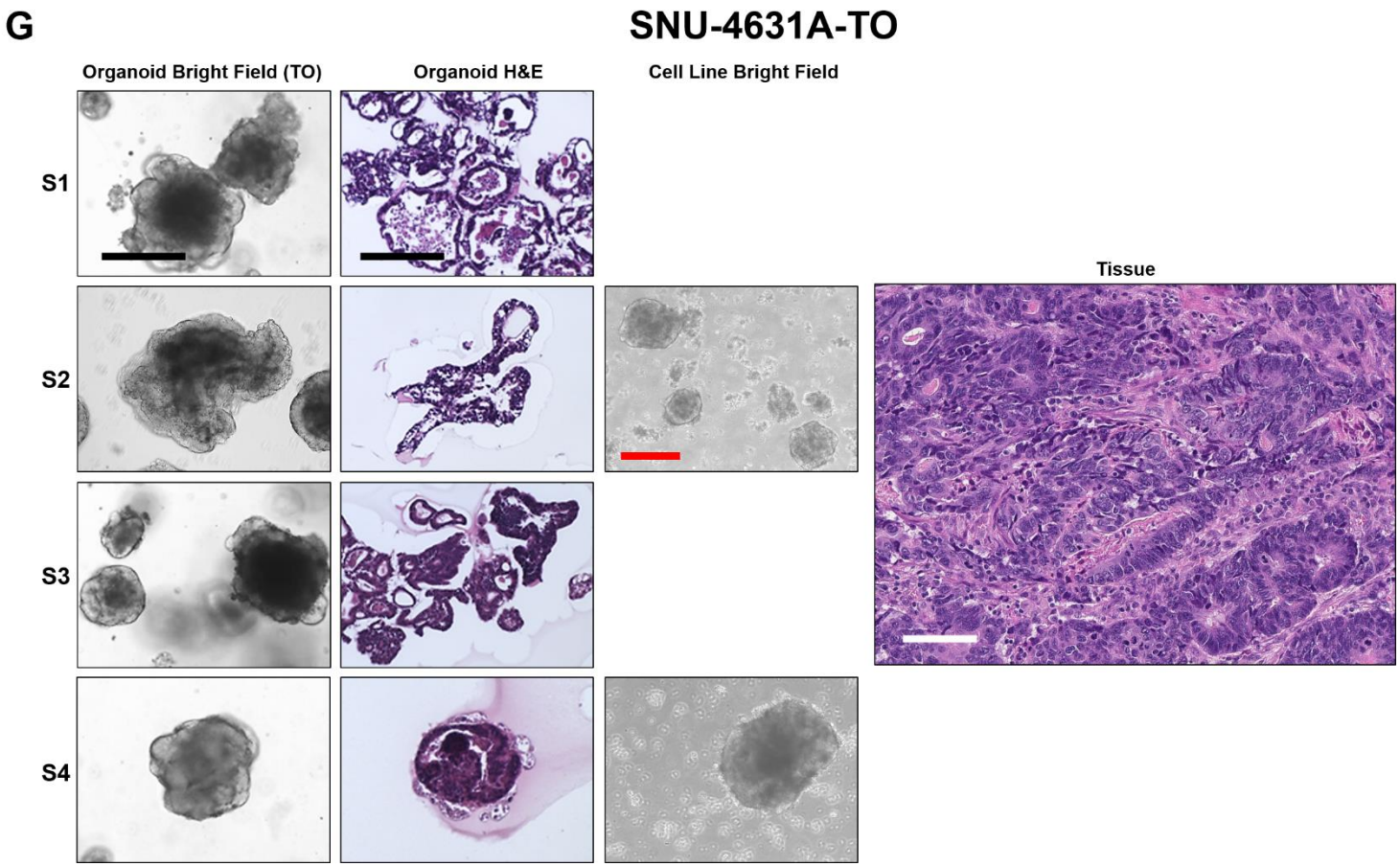

B
SNU-4146-TO
Organoid Bright Field (TO)
Organoid H&E
Cell Line Bright Field
Tissue
S1
S2
S3
S4

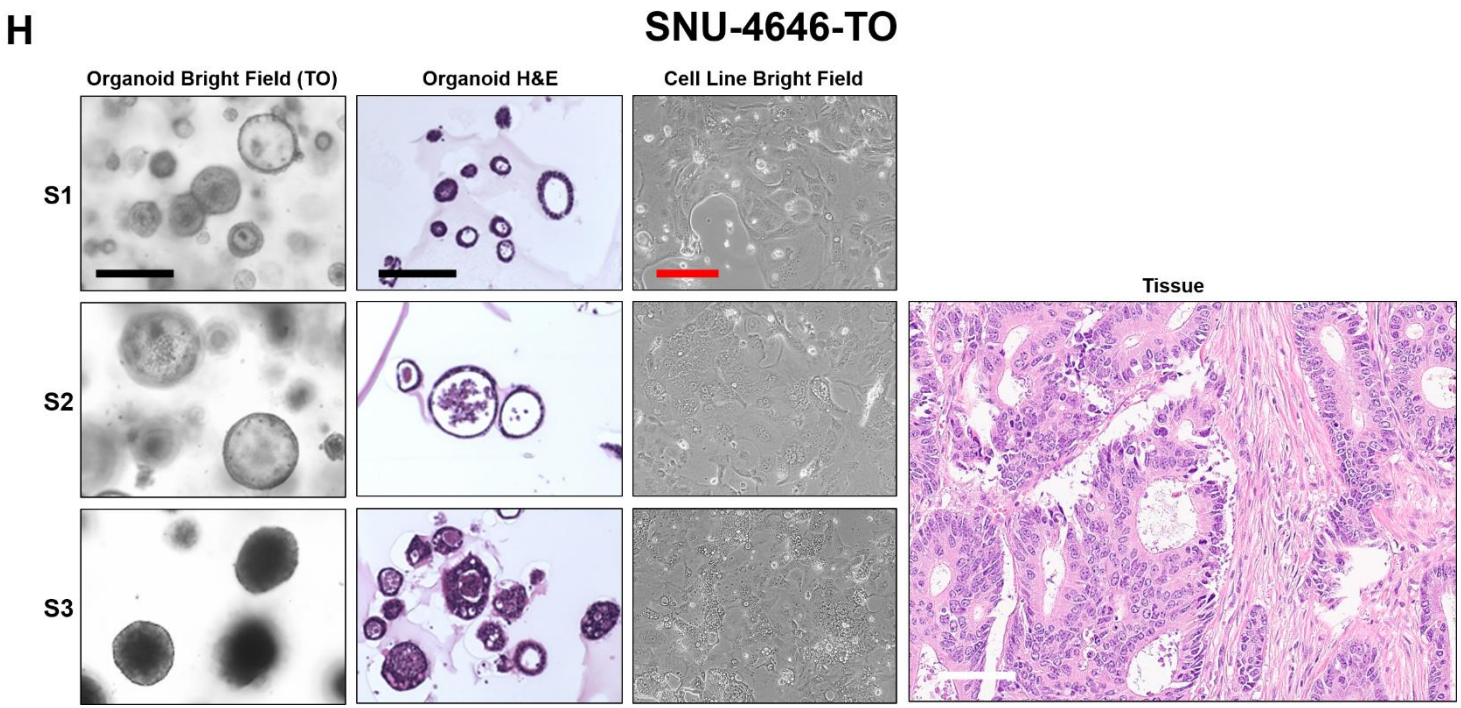

C
SNU-4351-TO
Organoid Bright Field (TO)
Organoid H&E
S1
S2
S3
S4
Tissue


D
SNU-4374-TO
Organoid Bright Field (TO)
Organoid H&E
Tissue
S1
S2
S3
S4

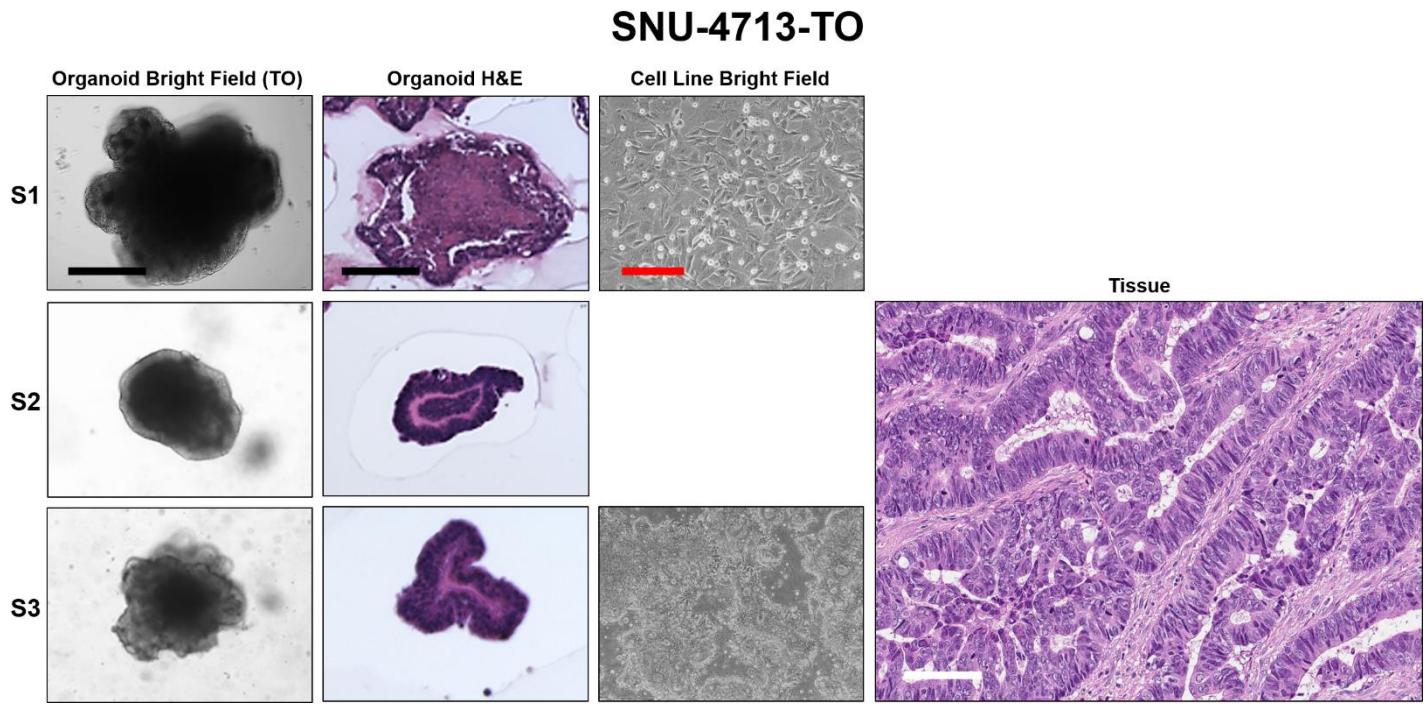

E
SNU-4376A-TO
Organoid Bright Field (TO)
Organoid H&E
Cell Line Bright Field
Tissue
S1
S2
S3
S4

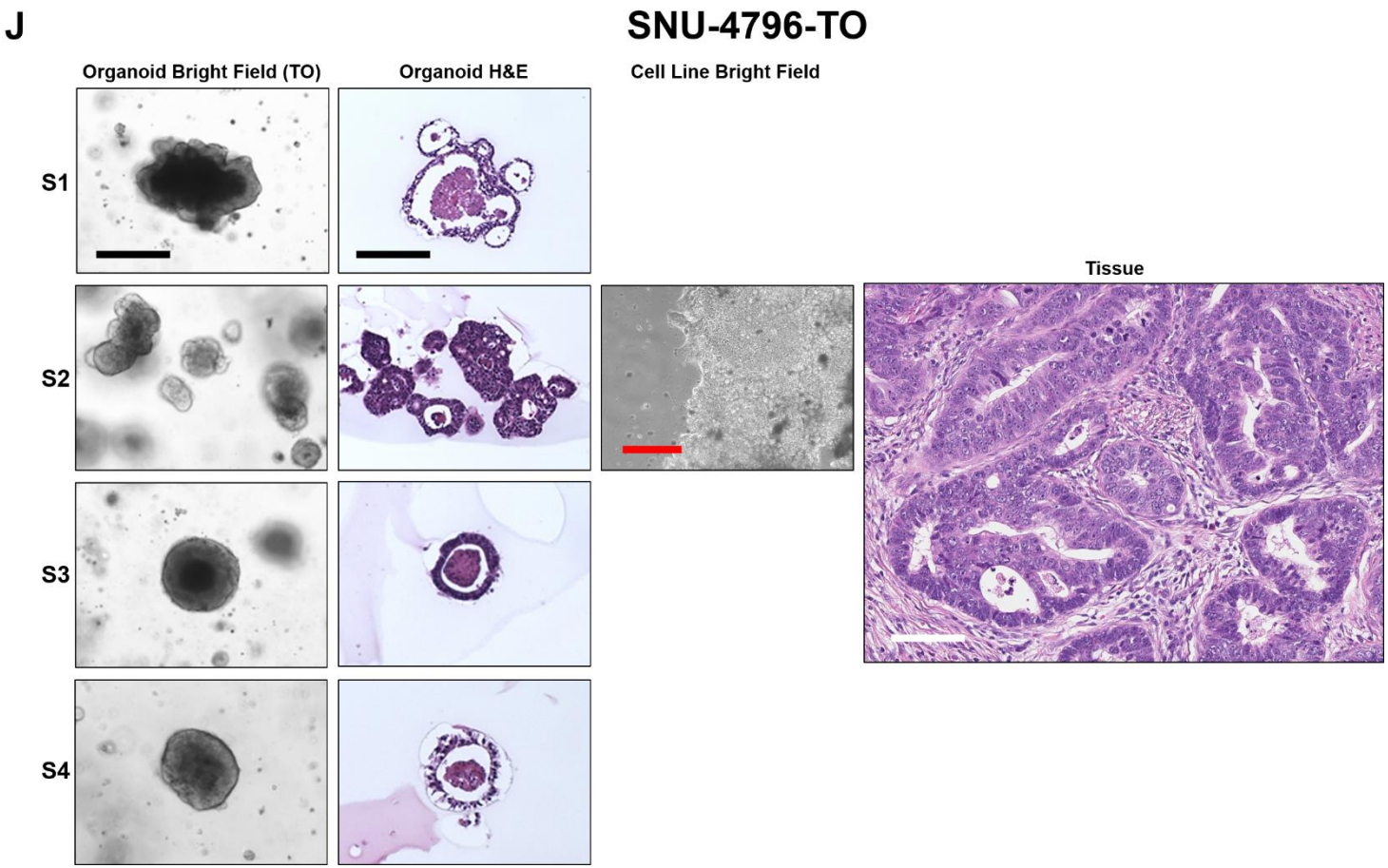

F
SNU-4398-TO
Organoid Bright Field (TO)
Organoid H&E
Cell Line Bright Field
Tissue
S1
S2
S3
S4

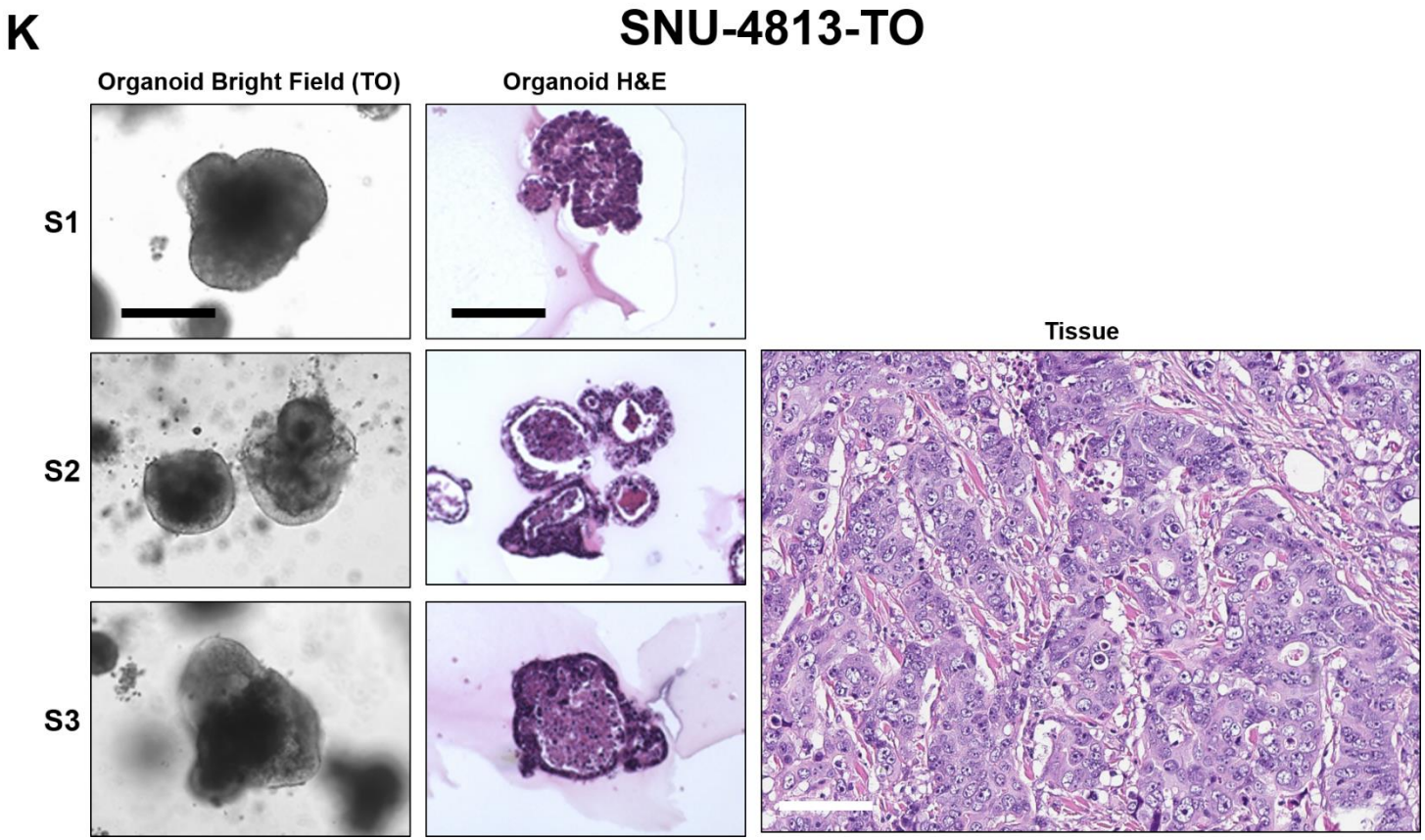

G
SNU-4631A-TO
Organoid Bright Field (TO)
Organoid H&E
Cell Line Bright Field
Tissue
S1
S2
S3
S4


H
SNU-4646-TO
Organoid Bright Field (TO)
Organoid H&E
Cell Line Bright Field
Tissue
S1
S2
S3

I


SNU-4713-TO
Organoid Bright Field (TO)
Organoid H&E
Cell Line Bright Field
Tissue
S1
S2
S3


J
SNU-4796-TO
Organoid Bright Field (TO)
Organoid H&E
Cell Line Bright Field
Tissue
S1
S2
S3
S4


K
SNU-4813-TO
Organoid Bright Field (TO)
Organoid H&E
S1
S2
S3
Tissue

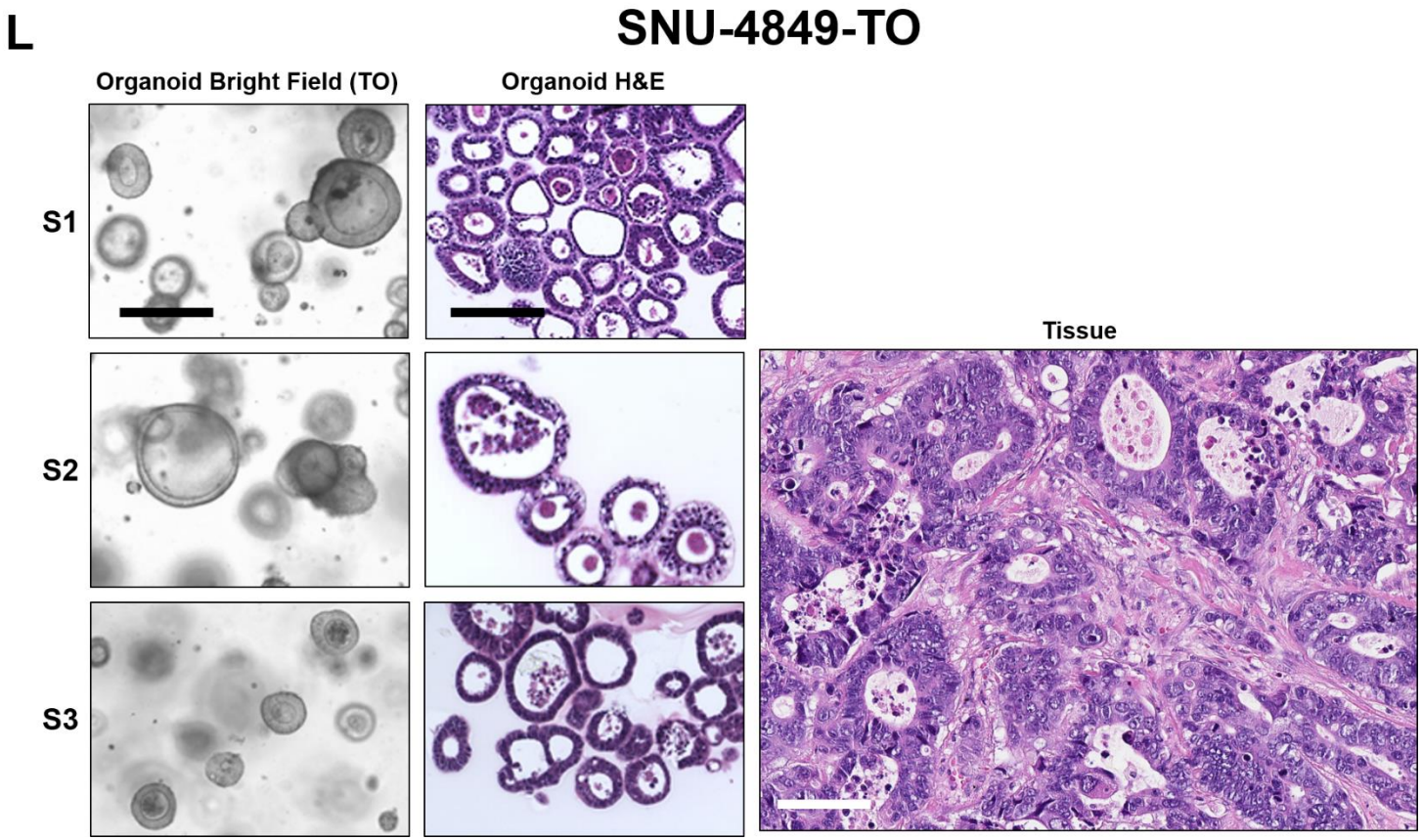
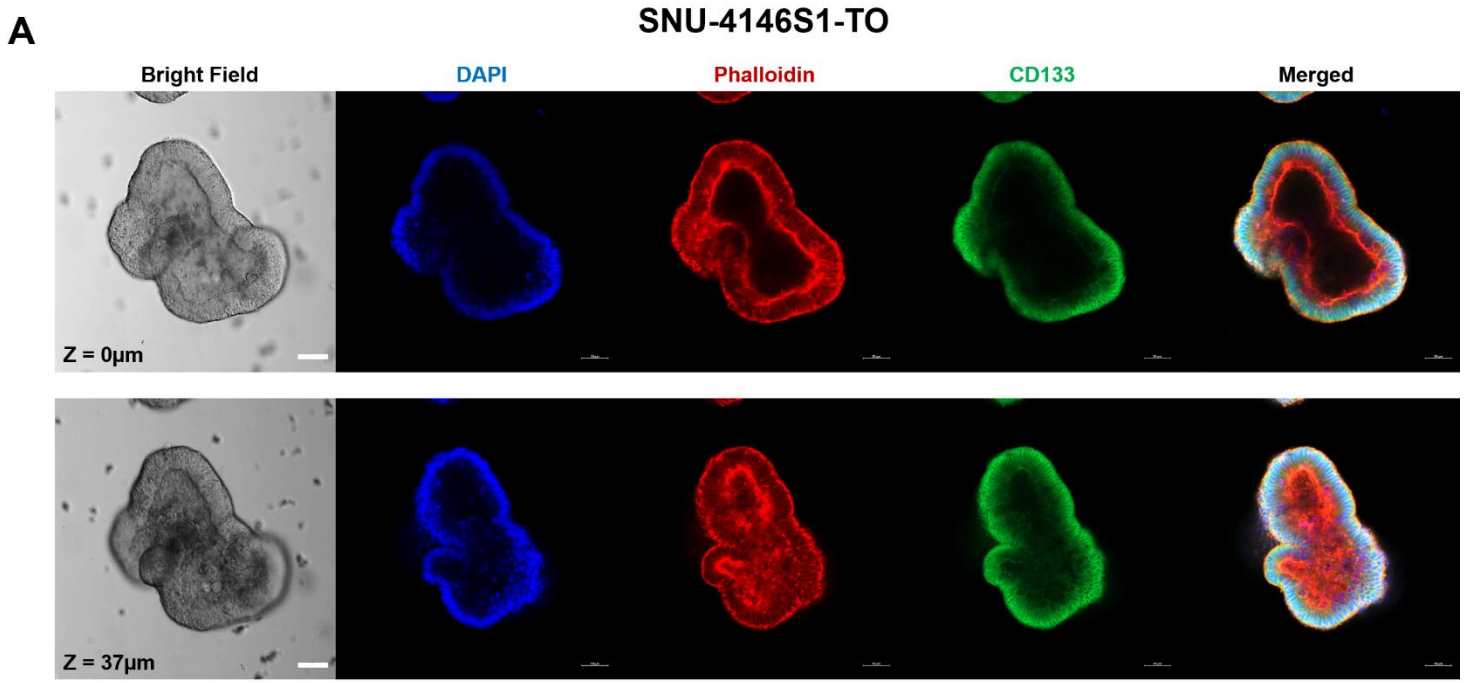
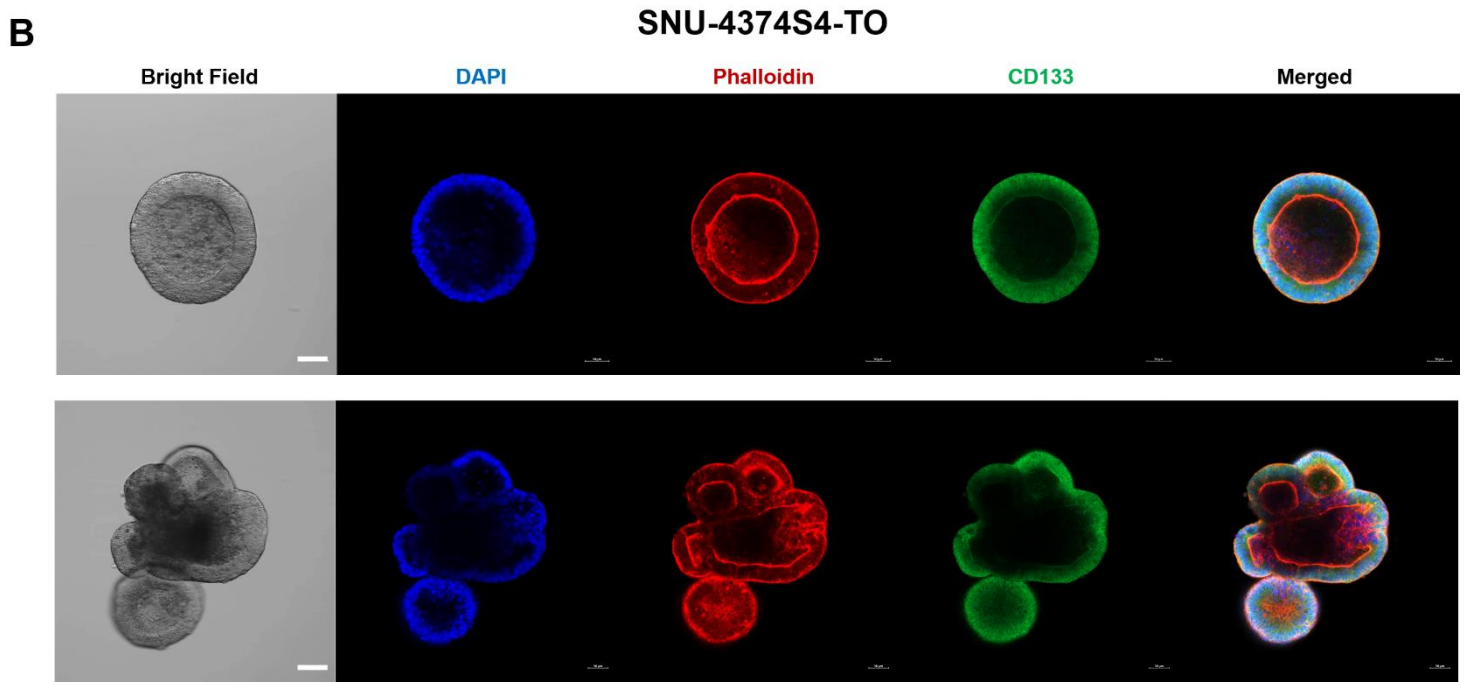

L
SNU-4849-TO
Organoid Bright Field (TO)
Organoid H&E
Tissue
S1
S2
S3

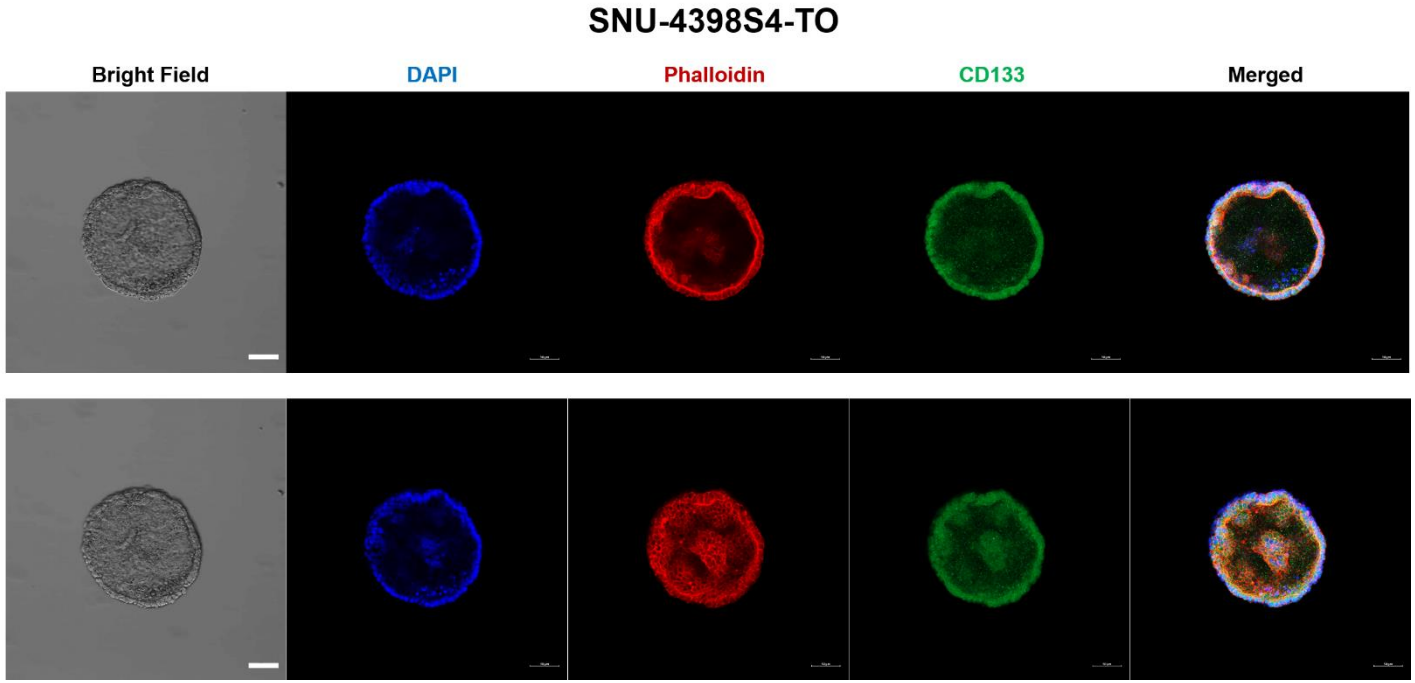
Figure 1A–L. Histopathological Characterization of Patients–Derived Cell lines and Organoids. A–L represent SNU–4139, SNU–4146, SNU–4351, SNU–4374, SNU–4376A, SNU–4398, SNU–4631A, SNU–4646, SNU–4713, SNU–4796, SNU–4813, and SNU–4948 series respectively.

Organoids architecture resembles primary tumor epithelium. H&E staining of primary tumor and the tumor organoids derived of these. A feature of most organoids is the presence of one or more lumens, resembling the tubular structures of the primary tumor. Tumors devoid of lumen give rise to compact organoids without lumen. In vitro cultivation, the organoid lines exhibited spheroidal, asymmetric and loose aggregates morphologies. The matched tumor cell lines grew as monolayers of substrate–adherent cells displaying mostly polygonal and spindle morphology. Hematoxylin–eosin (H&E) staining of formalin fixed paraffin embedded (FFPE) organoid sections outlined that tumor–derived organoids presented patient–explicit heterogeneous morphologies ranging from thin–walled cystic structures to solid/compact structures devoid of a lumen. Black scale bar = 250 μM, Red scale bar = 70 μM, White scale bar = 105 μM.


A
SNU-4146S1-TO
Bright Field
DAPI
Phalloidin
CD133
Merged
Z = 0μm
Z = 37μm


B
SNU-4374S4-TO
Bright Field
DAPI
Phalloidin
CD133
Merged

C



D

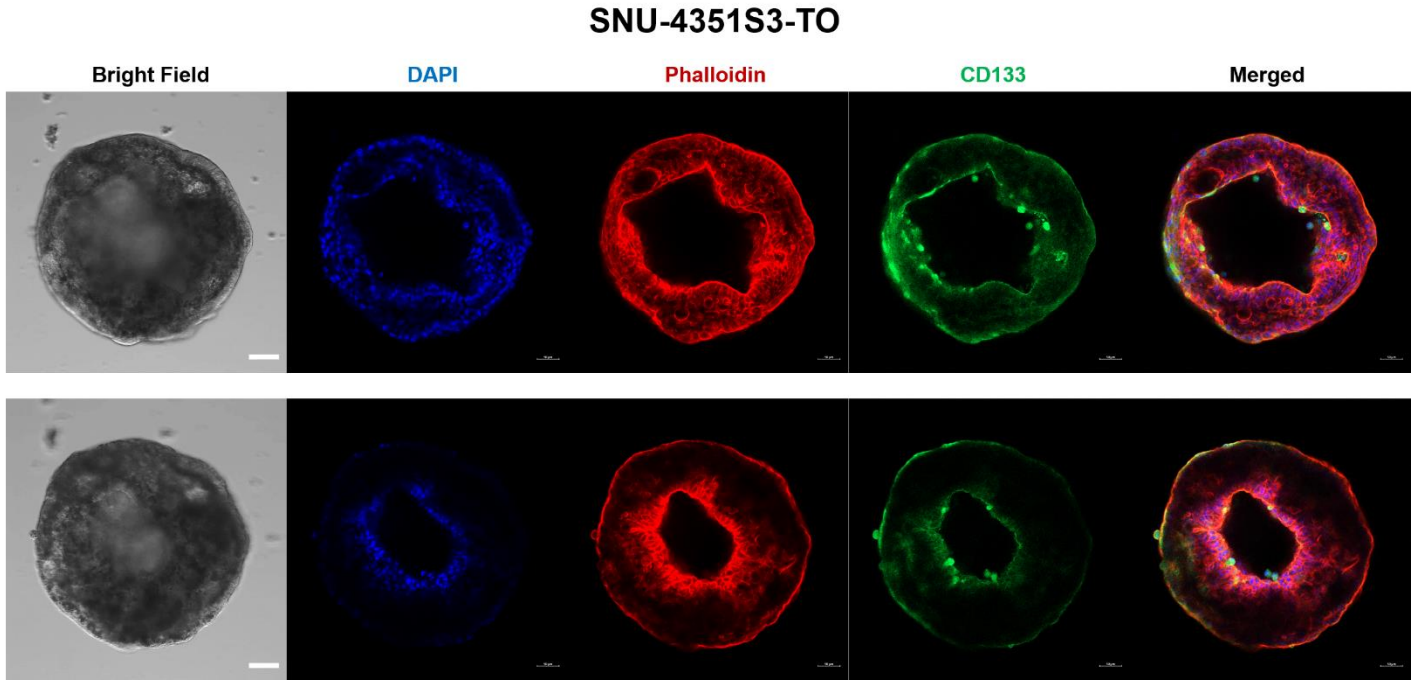
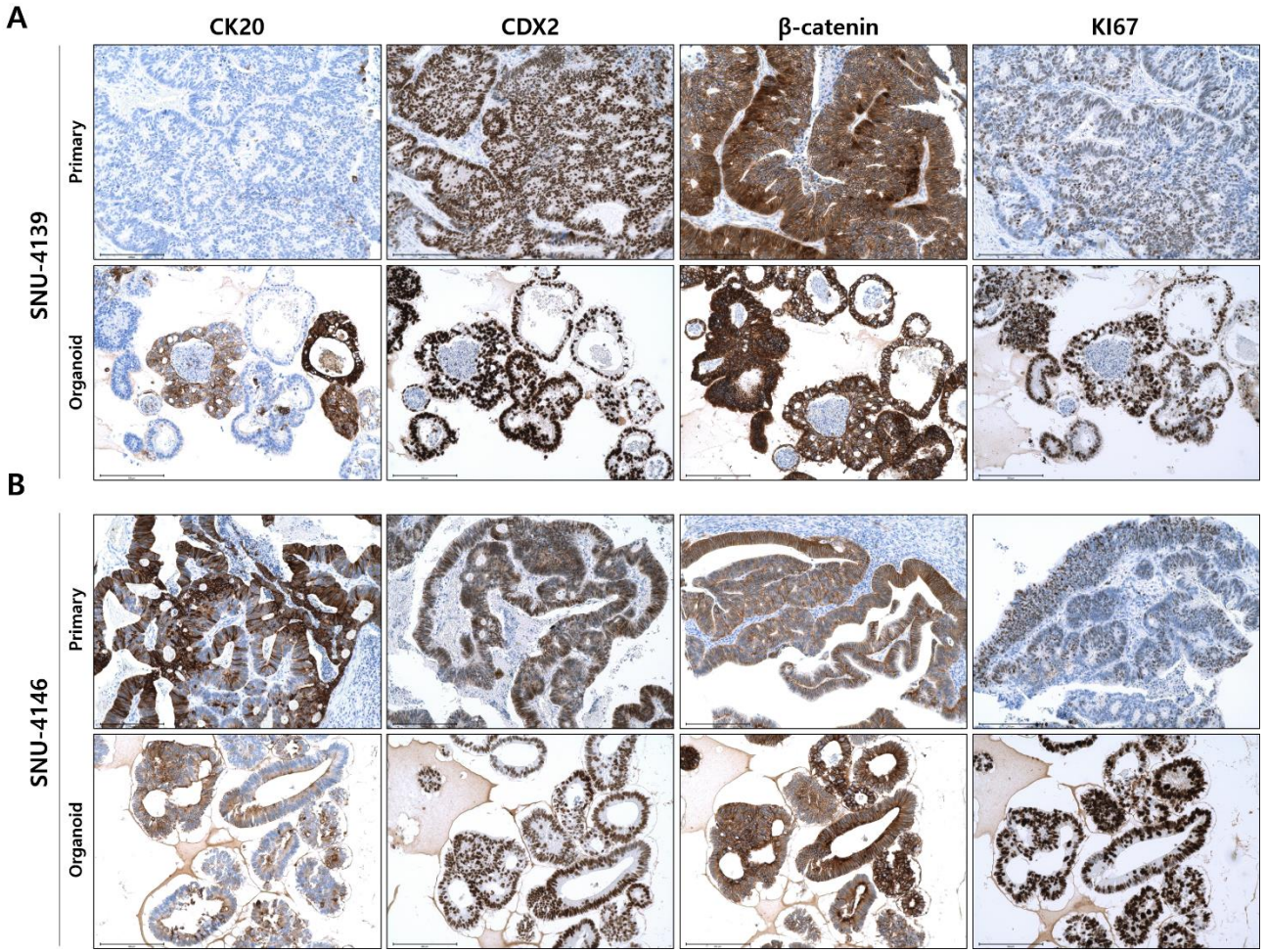
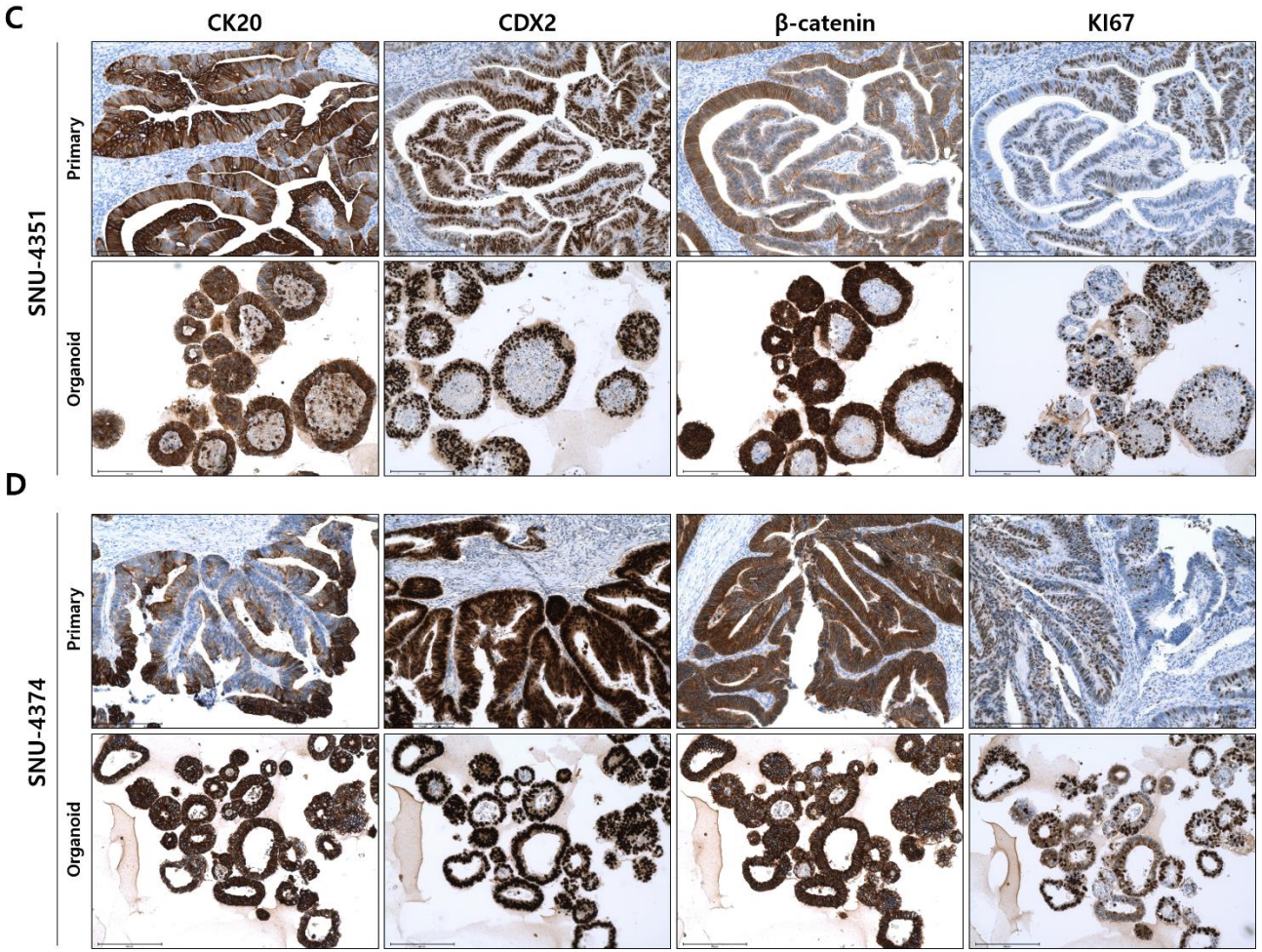
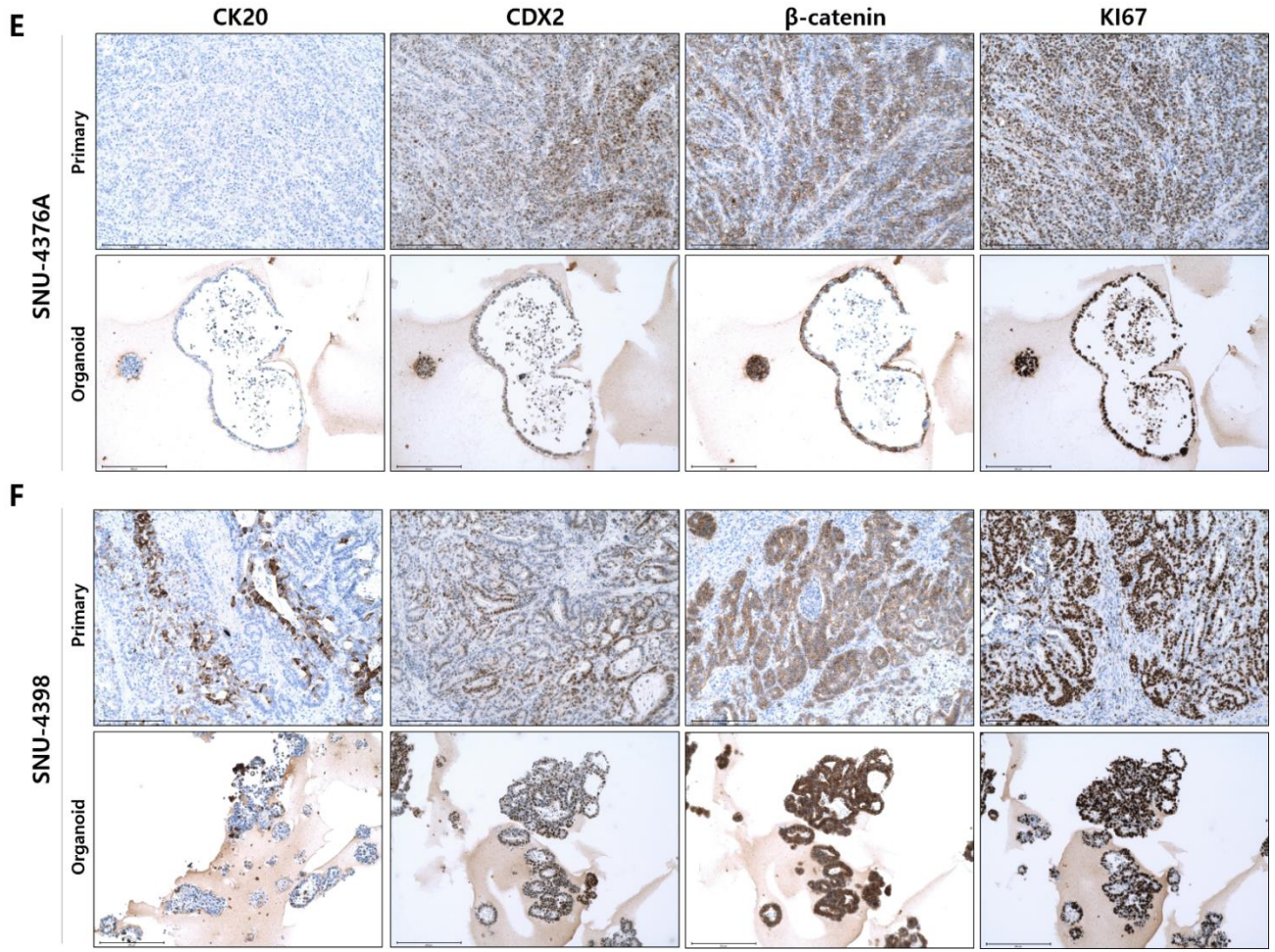


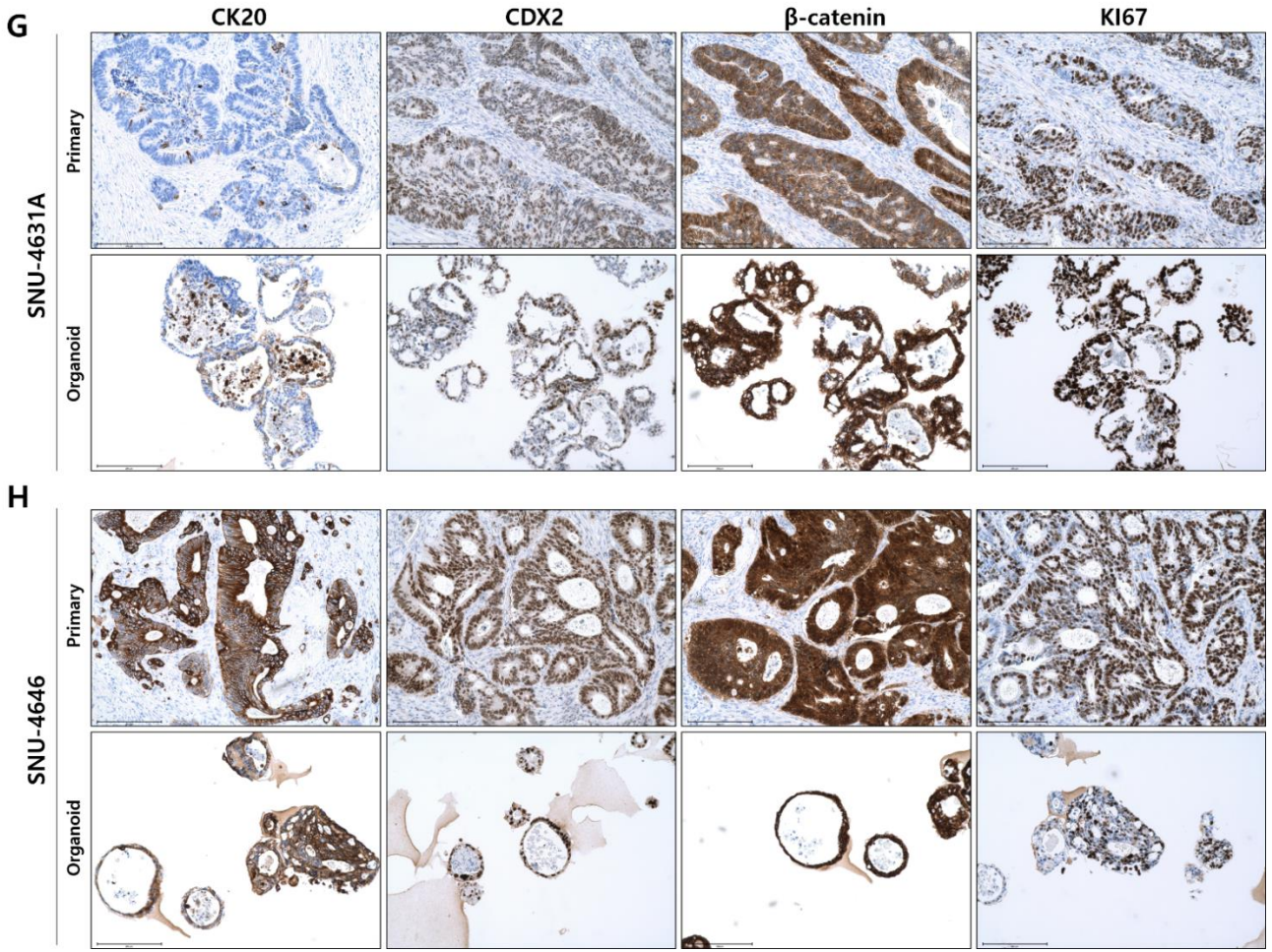
Figure 2A-D. Immunocytochemistry of Patients-Derived Organoids. A-D represent SNU-4146S1, SNU-4374S4-TO, SNU-4398S4-TO and SNU-4351S3-TO series respectively. Scale bar = 50 μM.

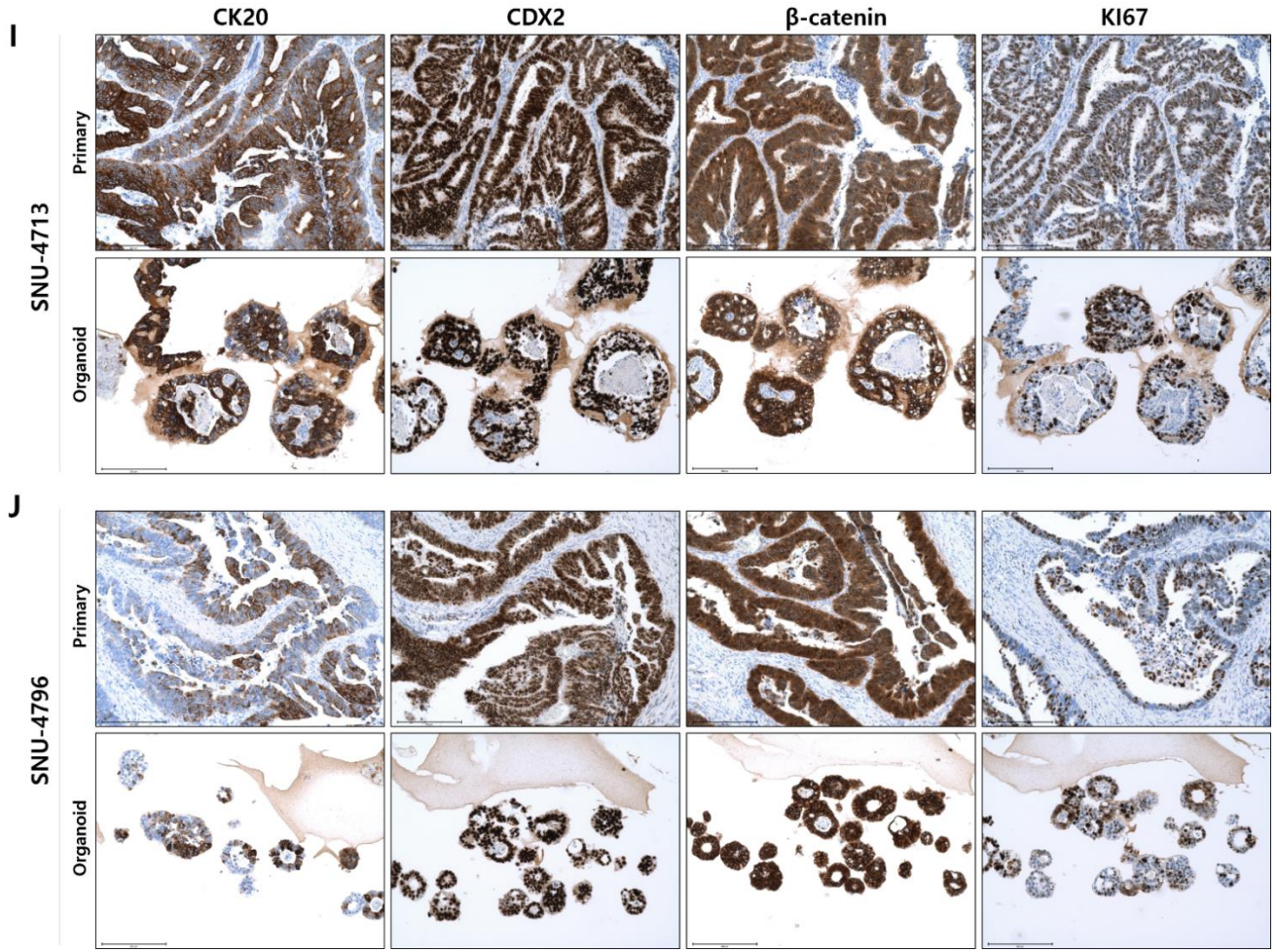
The intra- or inter-morphological variances of established organoids were further inspected with immunocytochemistry. Figure 2A indicated that SNU-4146S1-TO maintained the crypt-like structure of original colorectal tumor tissue. Figure 2B displayed that there existed intra-morphological heterogeneity in SNU-4374S4-TO. One subclone grew as spheroidal and asymmetric shape, whereas another subclone retained thin-walled cystic structures with a lumen, which was already confirmed in H&E staining (Figure 1D-S4). Figure 2C showed that the internal configuration of solid/compact organoid with a thin lumen consists of not only dirty necrosis clumps but also living cells with distinct actin structure. Prominent CD133 expressing cells were detected in inner luminal region of SNU-4351S3-TO, which may suggest that the organoids maintained the cancer stem cells.


A
CK20
CDX2
β-catenin
KI67
SNU-4139
Primary
Organoid
B
SNU-4146
Primary
Organoid


C
CK20
CDX2
β-catenin
KI67
SNU-4351
Primary
Organoid
D
SNU-4374
Primary
Organoid


E
CK20
CDX2
β-catenin
KI67
SNU-4376A
Primary
Organoid
F
SNU-4398
Primary
Organoid


G
CK20
CDX2
β-catenin
KI67
SNU-4631A
Primary
Organoid
H
SNU-4646
Primary
Organoid


I
CK20
CDX2
β-catenin
KI67
SNU-4713
Primary
Organoid
J
SNU-4796
Primary
Organoid

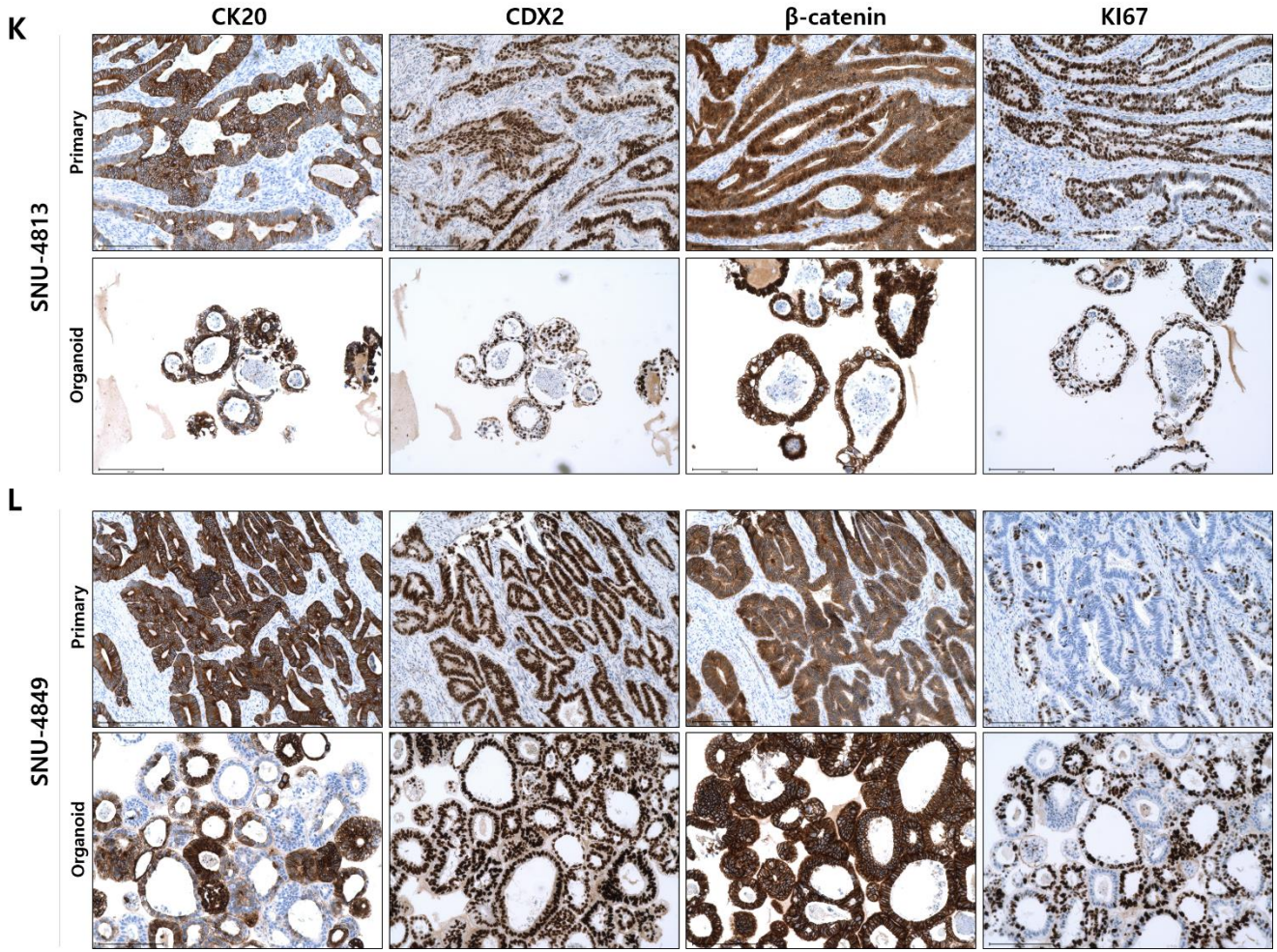
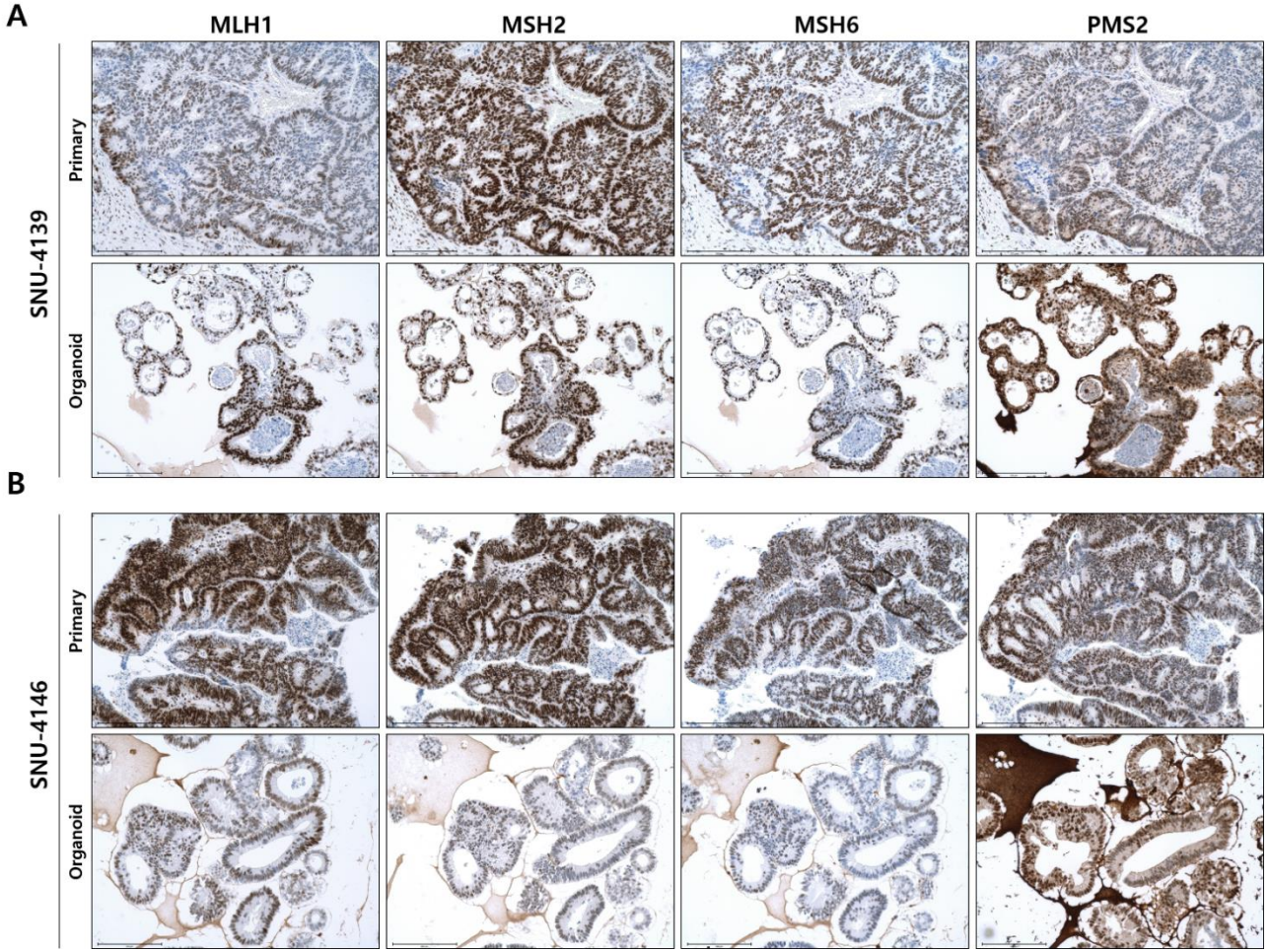
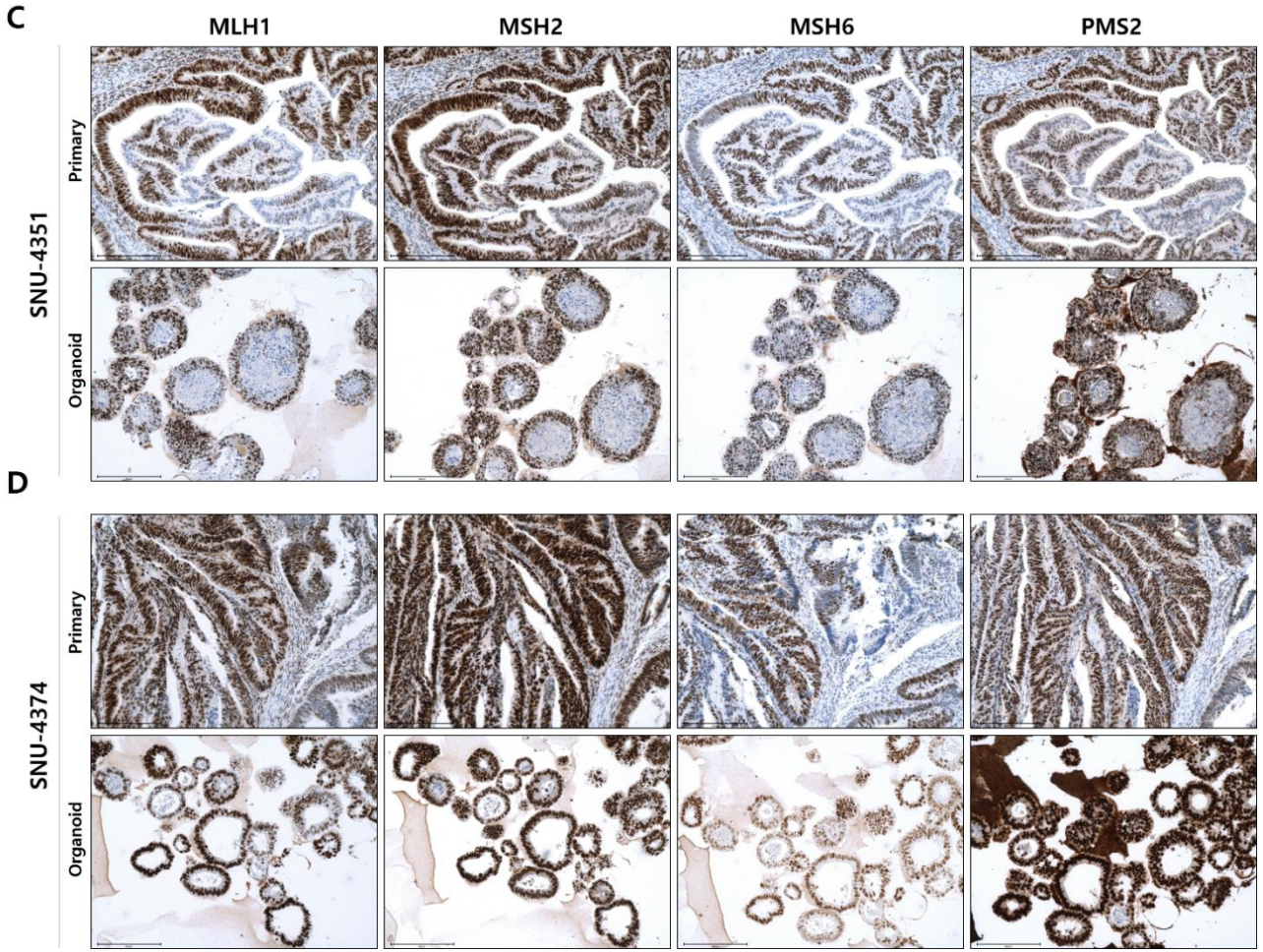
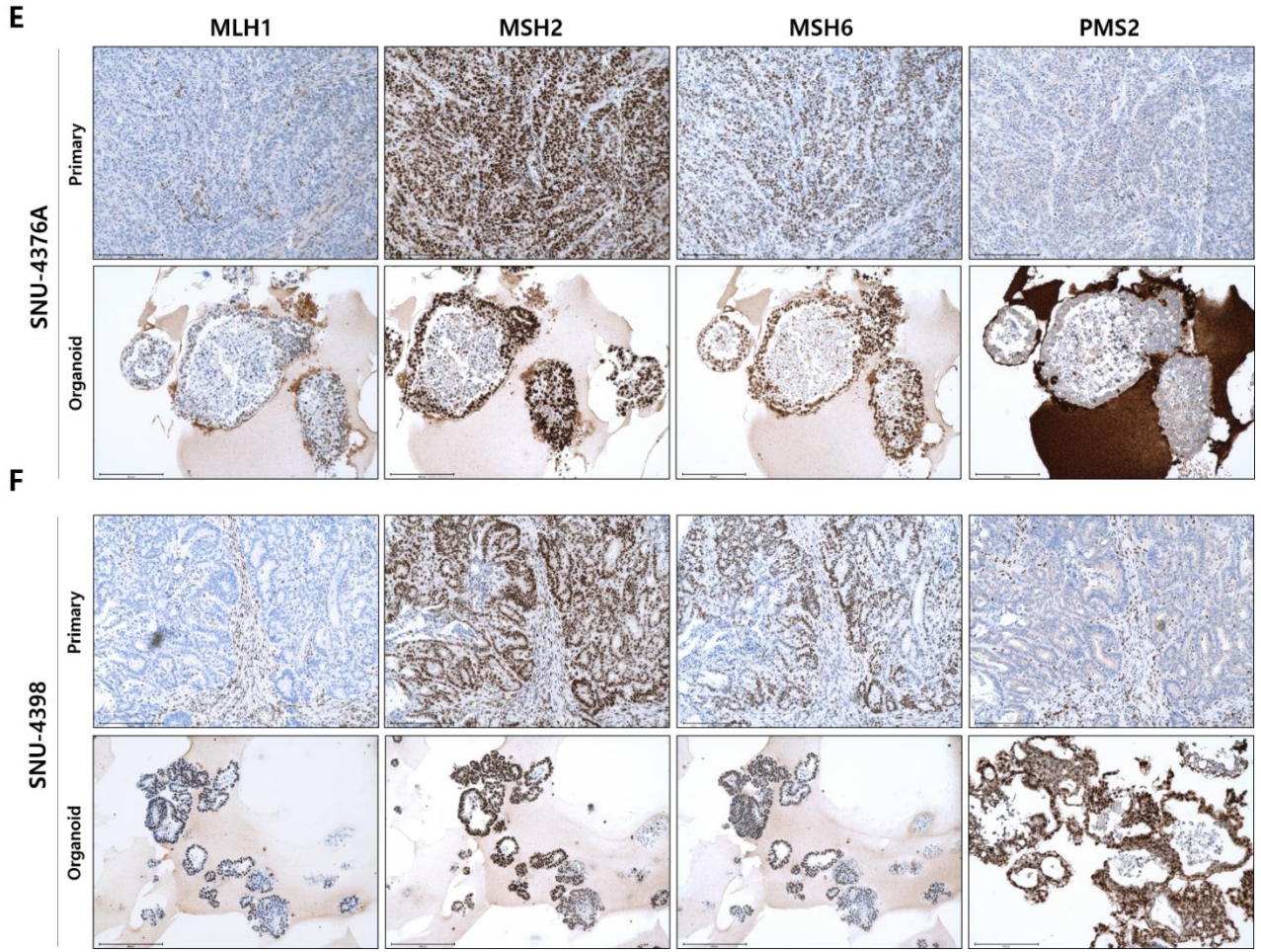

K
CK20
CDX2
β-catenin
KI67
SNU-4813
Primary
Organoid
L
SNU-4849
Primary
Organoid

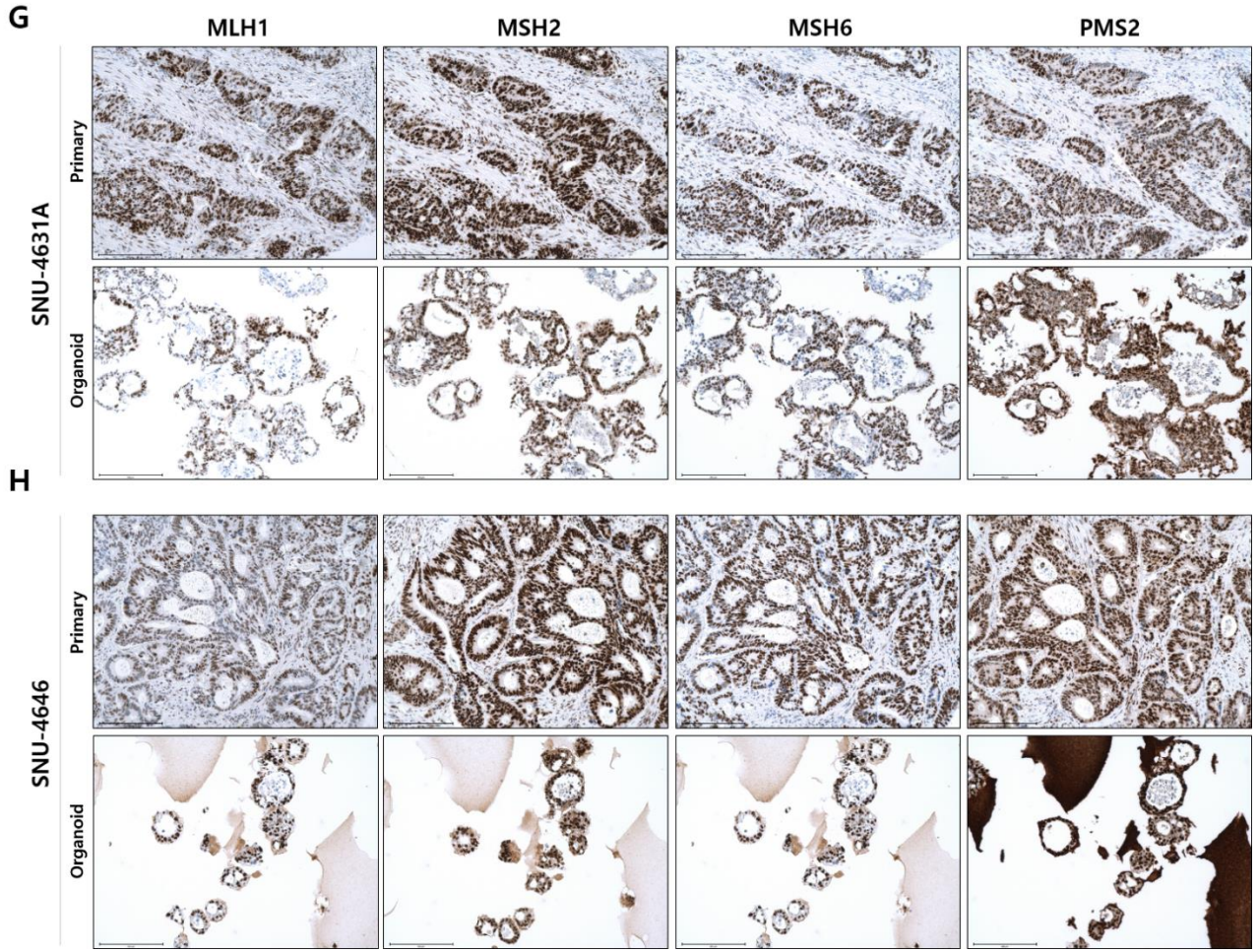
Figure 3A-L. Immunohistochemistry of Patients-Derived Organoids. A-L represent SNU-4139, SNU-4146, SNU-4351, SNU-4374, SNU-4376A, SNU-4398, SNU-4631A, SNU-4646, SNU-4713, SNU-4796, SNU-4813, and SNU-4948 series respectively. Scale bar = 200 μM.

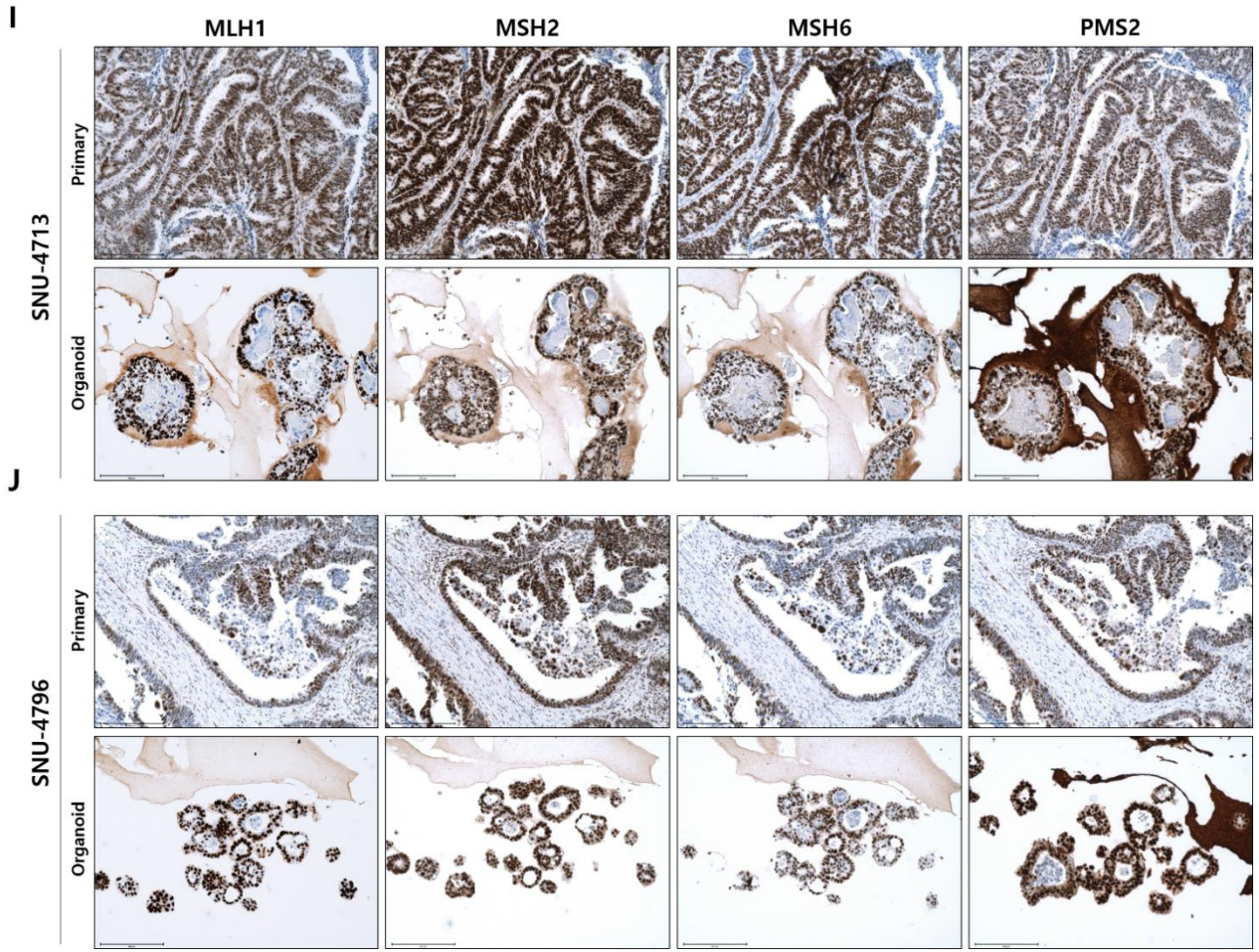
Cytokeratin 20 (CK20) and caudal type homeobox 2 (CDX2), as well as nuclear β-catenin and KI-67 were quantified and compared between matched organoids and patient tumors. Organoids retained similar presence and intensity of these markers.


A
MLH1
MSH2
MSH6
PMS2
SNU-4139
Primary
Organoid
B
SNU-4146
Primary
Organoid


C
MLH1
MSH2
MSH6
PMS2
SNU-4351
Primary
Organoid
D
SNU-4374
Primary
Organoid


E
MLH1
MSH2
MSH6
PMS2
SNU-4376A
Primary
Organoid
F
SNU-4398
Primary
Organoid


G
MLH1
MSH2
MSH6
PMS2
SNU-4631A
Primary
Organoid
H
SNU-4646
Primary
Organoid


I
MLH1
MSH2
MSH6
PMS2
SNU-4713
Primary
Organoid
J
SNU-4796
Primary
Organoid

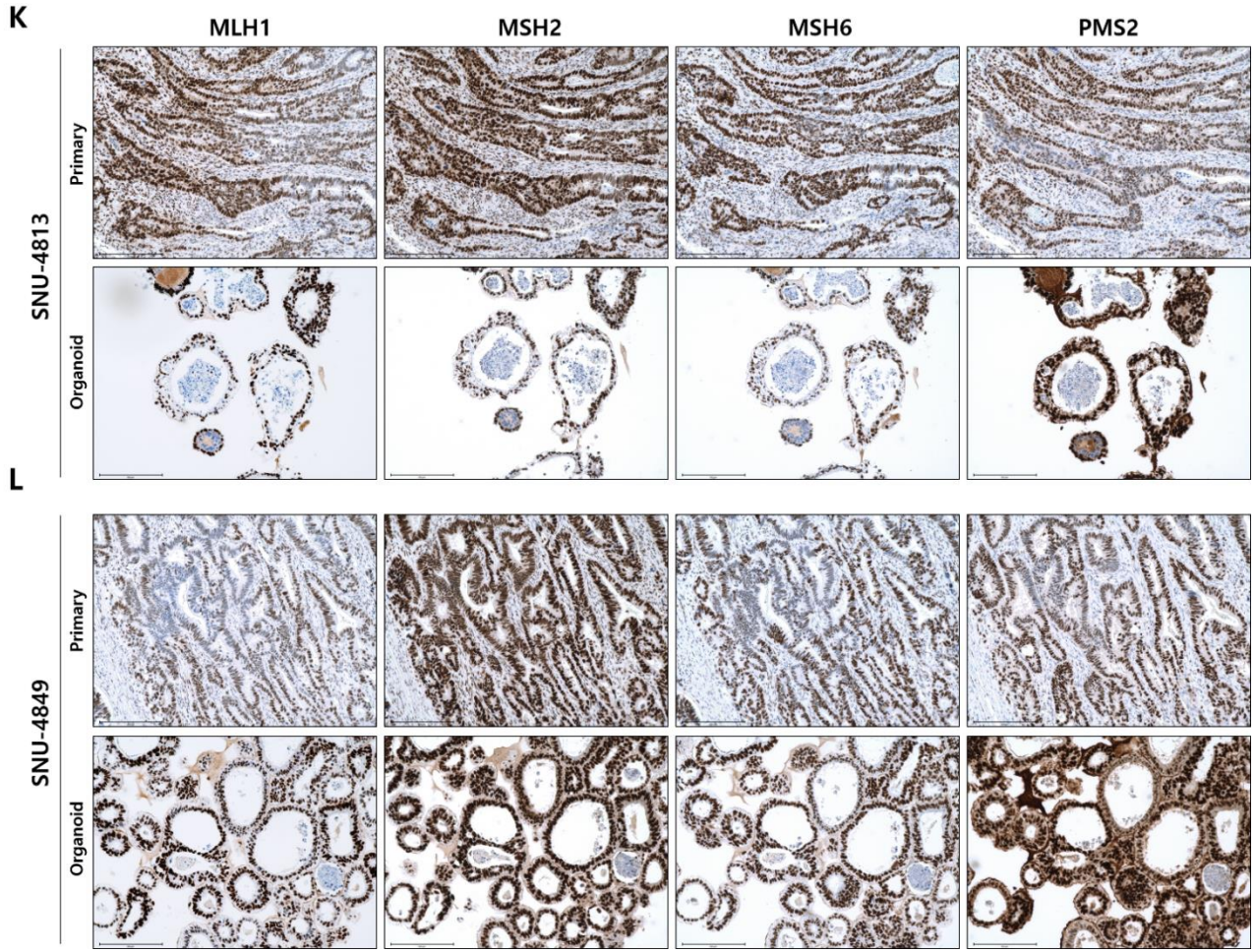

K
MLH1
MSH2
MSH6
PMS2
SNU-4813
Primary
Organoid
L
SNU-4849
Primary
Organoid

Figure 4A-L. Comparison of nuclear mismatch repair proteins between patient and organoid samples. A-L represent SNU-4139, SNU-4146, SNU-4351, SNU-4374, SNU-4376A, SNU-4398, SNU-4631A, SNU-4646, SNU-4713, SNU-4796, SNU-4813, and SNU-4948 series respectively. Scale bar = 200 μM.

MutL homolog 1 (MLH1), MutS Homolog 2 (MSH2), MutS Homolog 6 (MSH6), PMS1 Homolog 2 (PMS2) were quantified and compared between matched organoids and patient tumors. Organoids retained similar presence and intensity of these markers. SNU-4376A and SNU-4398 are deficient in MSH1 and PMS2.

Overall, these data validated the histological characters and expression pattern of specific markers are reiterated in the organoids.

Organoid Lines Recapitulate the Genomic Features of Human Colorectal Cancer

Many studies have shown that patient-derived organoids recapitulated the genomic landscapes of original tumors, including mutations and copy number variations (CNVs) (35-37). To determine genomic concordance between parental tumors and patient-derived cell lines/organoids and illustrate the evolutionary trajectories in the colorectal tumorigenesis, we performed whole-exome sequencing (WES) on twelve colorectal tumor cases. For each case, we established three to four multiregion-derived 2D cell lines and/or organoids, and sequenced a matched tumor and normal mucosa sample as a control, which amounted to 23 cancer cell lines, 42 organoids, 11 tumor samples, and 12 normal samples in total. The normal/tumor tissue for patient 4376 were not available, and normal tissue was alternated by matched blood DNA for somatic variants calling and CNV analysis. The result of WES is summarized in Table

3. WES identified multiple genomic mutations in the applied samples including point mutations in putative tumor driver genes, as well as copy number variations. Among the 42 tumor organoids and 23 tumor cell lines, 7 organoids and 7 cell lines which were originated from patient 4376A and 4398 exhibited hyper-mutation (>10 mutations/Mb) (Figure 5). The percentage of hypermutated organoids in the patient panel was 16.6%; 2 of 12, which was accorded with the described frequency in a larger cohort of clinical samples (19). The most predominant point mutation type was C to T transitions at CpG (Figure 6A-L, Table 5), in parallel with other large cohort CRC sequencing (19). Mutational signature analysis indicated that cell lines and organoids that were originated from a same patient displayed highly concordant pattern (Figure 6A-L, Figure 7, Table 6). Mutational concordance within the coding regions in both tumor cell lines and organoids was highly corresponded with the matched tumor specimen for both hypermutated and non-hypermutated patients (Figure 8) (median = 0.90 frequency of concordance, range 0.87 to 0.94). Cell line/Organoid-specific and tumor-specific discordant alterations were analyzed for their genetic

significance in tumorigenesis based on data reported from the PanCancer analysis of 10,000 TCGA tumor samples (31). On average, 5.7% (43/744) of discordant mutations found in cell line/organoid-specific affected cancer-driver genes, counting a third hit to APC (c.3334dup/p.T1112Nfs*7) with variant allele frequency (VAF) of 0.95 , TP53 (c.349_355dup/p.A119Gfs*32) with VAF of 0.95, SMAD2 (c.341G>A/p.R114H) with VAF of 0.14, and CDH1 (c.208del/p.S70Pfs*13) with VAF of 0.14. On average, 3.5% (29/831) of discordant mutations detected in tumor tissue-specific represent cancer-driver genes, including APC (c.7749del/p.A2584Qfs*9) with 0.47 variant allele frequency (VAF), TP53 (c.91G>A/p.V31I) with 0.25 VAF, PIK3CA (c.320A>G/p.N107S) with 0.42 VAF, MSH3 (c.181_189dup/p.A61_P63dup) with 0.44 VAF. The discordant mutations had a mean allelic frequency of 37.2% and 29.9% for the tumor tissue and organoids, respectively. The relatively low allelic frequency of the discordant mutations could suggest gained mutations during proliferation or derivation, as well as the diminution or enrichment of a sub-clonal population in the organoid cultivation present within the original tumor tissue. The most frequently mutated

genes in CRC (31, 38, 39) were recapitulated in the cell line/organoid cultures. Inactivating mutations to the tumor suppressors TP53, APC, and FBXW7 as well as activating alterations in KRAS (codon 12) and PIK3CA (codon 107, 542, 545, 939, 1044 and 1047) were observed. Activating mutations in BRAF and TGFBR1/2 mutations were not observed in our cohort (Figure 6, Table 4).

Genetic alternations in DNA mismatch repair (MMR)-associated pathways are concomitant with a hypermutation (40). Missense mutations were present in MSH3 in SNU-4398, and POLE mutations were detected in both SNU-4376A and SNU-4398 in accordance with their classification as hypermutated CRC cases (41). Interestingly, solely SNU-4398S2 acquired unique pathogenic frameshift mutation in MSH3 (c.1148del/p.K383Rfs*32) among SNU-4398 series. Besides, SNU-4146 patient tissue harbored missense mutation in MSH3 (c.82T>G/p.F28V) which was not transferred to any of its derivate. Majority of CRC cases harbor activating mutations in CTNNB1 or inactivation mutations in APC, AXIN2, FBXW7 and FAM123B (41). We found APC alterations in all but 2 of the series (SNU-4631A and SNU-4796). Neither of the

series carried activating mutations in CTNNB1. SNU-4631A had missense mutation in FBXW7. Epigenetic regulation takes significant parts in initiation and progression of CRC (42). In our cohort, various epigenetic factors such as KMT2C, KMT2D, ARID1A, and KDM6A, which are commonly detected in CRC (19) were mutated at a high rate. We also detected less-frequent driver alterations such as mutations in STAG2, SMA7, and APC2. Overall, the mutational spectrums identified in our cell lines/organoids reflect genomic features of their parental tumor.

Table 3. Whole Exome Sequencing Technical Information

SAMPLECODE	Mean Depth of Target Regions	GC(%)	Q20(%)	Q30(%)
SNU-4139_CT	101.6	50.1	98.2	95.2
SNU-4139_NT	87.3	50.7	97.9	95.5
SNU-4139S1	90.4	51.9	98.1	94.8
SNU-4139S1-TO	107.6	51.1	98.5	94.8
SNU-4139S2-TO	101.5	51.5	98.4	94.8
SNU-4139S3	102.2	50.5	98.0	94.6
SNU-4139S3-TO	97.3	51.2	98.3	95.3
SNU-4139S4-TO	81.3	50.0	97.9	95.3
SNU-4146_CT	100.1	51.6	97.5	95.4
SNU-4146_NT	102.2	51.1	98.4	95.1
SNU-4146S1T	103.9	51.3	98.0	95.1
SNU-4146S1-TO	102.8	51.8	98.3	94.8
SNU-4146S2	107.0	51.4	98.2	95.0
SNU-4146S2-TO	91.3	50.1	97.8	94.9
SNU-4146S3	86.0	51.6	98.0	94.4
SNU-4146S3-TO	81.2	50.8	98.2	94.9
SNU-4146S4	92.3	50.8	98.2	95.4
SNU-4146S4T	87.8	51.2	98.2	95.0
SNU-4146S4-TO	90.9	50.4	97.9	95.3
SNU-4351_CT	98.6	50.6	98.2	95.6
SNU-4351_NT	95.1	50.7	97.4	93.5
SNU-4351S1-TO	89.0	50.7	97.8	93.9
SNU-4351S2-TO	107.4	50.6	97.9	95.3
SNU-4351S3-TO	84.2	51.6	98.2	95.2
SNU-4351S4-TO	99.0	50.4	97.8	95.4
SNU-4374_CT	99.1	50.2	97.5	94.2
SNU-4374_NT	95.6	50.4	97.6	95.2
SNU-4374S1-TO	84.8	50.6	98.3	95.5
SNU-4374S2-TO	93.0	51.1	98.1	94.8
SNU-4374S3-TO	102.2	51.9	98.1	94.9
SNU-4374S4-TO	103.9	51.3	97.9	94.7
SNU-4376_CT	100.9	52.6	98.8	95.3
SNU-4376_NT	95.5	50.1	99.5	95.1
SNU-4376AS1	91.2	50.8	98.3	94.7
SNU-4376AS1T	105.0	51.3	98.0	94.5
SNU-4376AS1-TO	91.9	51.3	98.4	95.4
SNU-4376AS2	107.8	51.0	98.5	95.4
SNU-4376AS3	94.6	51.0	97.9	94.6
SNU-4376AS3T	106.5	51.5	98.2	95.0
SNU-4376AS3-TO	95.6	51.1	98.0	95.2
SNU-4376AS4	90.0	51.9	97.9	94.9
SNU-4376AS4-TO	91.8	50.8	98.2	94.9
SNU-4398_CT	101.2	50.4	97.8	95.5
SNU-4398_NT	95.1	50.7	98.9	95.4
SNU-4398S1	89.0	51.0	98.0	94.6
SNU-4398S1-TO	100.4	52.0	98.4	95.0
SNU-4398S2	95.0	51.4	98.2	94.8
SNU-4398S2-TO	102.7	50.9	98.2	94.7
SNU-4398S3-TO	93.7	51.0	98.3	94.9
SNU-4398S4	91.4	51.5	98.1	94.6
SNU-4398S4-TO	92.0	50.3	97.9	95.5

Continued

SAMPLECODE	Mean Depth of Target Regions	GC(%)	Q20(%)	Q30(%)
SNU-4631A_CT	101.5	50.1	97.6	95.5
SNU-4631A_NT	88.5	50.8	97.4	95.4
SNU-4631AS1	93.3	52.0	98.1	94.7
SNU-4631AS1-TO	85.6	51.1	98.4	94.9
SNU-4631AS2-TO	103.0	51.2	98.0	95.0
SNU-4631AS3-TO	85.9	50.1	98.1	94.8
SNU-4631AS4	91.2	52.0	98.3	95.1
SNU-4631AS4T	92.6	51.7	98.5	94.8
SNU-4631AS4-TO	99.3	51.0	98.2	95.2
SNU-4646_CT	85.4	50.1	97.2	95.2
SNU-4646_NT	85.6	50.2	97.3	95.7
SNU-4646S1T	90.1	50.3	98.2	95.3
SNU-4646S1-TO	81.4	50.2	97.9	94.8
SNU-4646S2T	88.8	51.9	98.1	94.6
SNU-4646S2-TO	98.5	51.2	97.9	95.4
SNU-4646S3T	82.5	51.2	97.9	95.1
SNU-4646S3-TO	82.3	50.6	98.0	94.9
SNU-4713_CT	89.6	50.2	97.5	95.2
SNU-4713_NT	85.5	50.6	97.6	95.3
SNU-4713S1	96.6	51.5	98.2	94.9
SNU-4713S1T	108.4	50.9	98.2	94.8
SNU-4713S1-TO	89.6	50.0	98.0	94.9
SNU-4713S2-TO	87.9	51.8	98.0	94.6
SNU-4713S3	89.4	51.4	98.1	94.7
SNU-4713S3-TO	70.2	51.2	98.1	94.8
SNU-4796_CT	89.4	50.2	97.8	95.2
SNU-4796_NT	55.6	50.1	97.4	95.2
SNU-4796S1-TO	72.2	51.1	98.0	94.5
SNU-4796S2	92.7	51.3	98.0	94.5
SNU-4796S2-TO	75.2	51.3	97.9	94.3
SNU-4796S3-TO	87.7	50.6	97.9	95.5
SNU-4796S4-TO	73.8	51.4	97.9	94.3
SNU-4813_CT	101.2	50.8	97.6	94.7
SNU-4813_NT	88.5	50.9	97.2	94.5
SNU-4813S1-TO	83.2	51.5	98.1	94.7
SNU-4813S2-TO	82.6	51.5	98.2	95.0
SNU-4813S3-TO	77.1	51.1	98.1	94.7
SNU-4849_CT	89.5	50.9	97.5	94.7
SNU-4849_NT	89.4	50.9	97.6	94.6
SNU-4849S1-TO	89.9	51.5	98.1	94.7
SNU-4849S2-TO	82.9	51.6	98.3	94.7
SNU-4849S3-TO	87.5	51.6	98.1	94.7

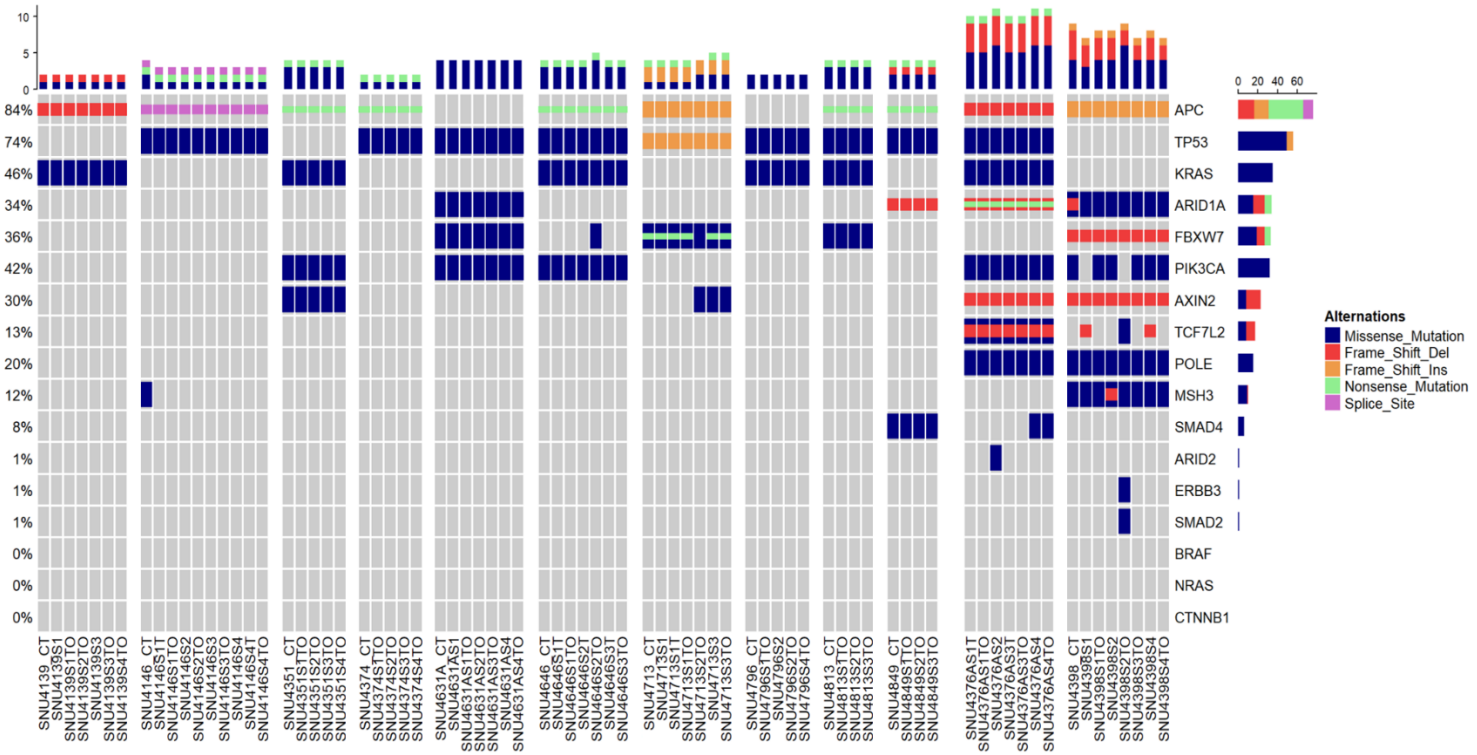

APC
TP53
KRAS
ARID1A
FBXW7
PIK3CA
AXIN2
TCF7L2
POLE
MSH3
SMAD4
ARID2
ERBB3
SMAD2
BRAF
NRAS
CTNNB1
Alternations
Missense_Mutation
Frame_Shift_Del
Frame_Shift_Ins
Nonsense_Mutation
Splice_Site

Figure 5. Heatmap of gene mutation variations in the most frequently mutated genes of colorectal cancer. WES identifies multiple genomic mutations in the twelve CRC series which amounted 23 cancer cell lines, 42 organoids, 11 tumor samples in total including point mutations in putative tumor driver genes. The most frequently mutated genes in CRC were recapitulated in the cell line/organoid cultures. Inactivating alterations to the tumor suppressors APC, TP53, and FBXW7 as well as activating mutations in KRAS (codon 12) and PIK3CA (codon 107, 542, 545, 939, 1044 and 1047) were observed. Activating mutations in BRAF and TGFBR1/2 mutations were not observed in our cohort. Among the 42 tumor organoids and 23 tumor cell lines, 7 organoids and 7 cell lines which were originated from patient 4376A and 4398 exhibited hyper-mutation (>10 mutations/Mb).

Table 4. Mutational profiles of twelve CRC series

SampleCode	Hugo_Symbol	Chromosome	HGVSc	HGVSp_Short	VAF
SNU-4139_CT	APC	chr5	c.4179del	p.D1394Ifs*21	0.95
SNU-4139_CT	KRAS	chr12	c.35G>T	p.G12V	0.33
SNU-4139S1-TO	APC	chr5	c.4179del	p.D1394Ifs*21	1.00
SNU-4139S1-TO	KRAS	chr12	c.35G>T	p.G12V	0.51
SNU-4139S1	APC	chr5	c.4179del	p.D1394Ifs*21	1.00
SNU-4139S1	KRAS	chr12	c.35G>T	p.G12V	0.36
SNU-4139S2-TO	APC	chr5	c.4179del	p.D1394Ifs*21	0.98
SNU-4139S2-TO	KRAS	chr12	c.35G>T	p.G12V	0.35
SNU-4139S3-TO	APC	chr5	c.4179del	p.D1394Ifs*21	1.00
SNU-4139S3-TO	KRAS	chr12	c.35G>T	p.G12V	0.43
SNU-4139S3	APC	chr5	c.4179del	p.D1394Ifs*21	0.98
SNU-4139S3	KRAS	chr12	c.35G>T	p.G12V	0.60
SNU-4139S4-TO	APC	chr5	c.4179del	p.D1394Ifs*21	1.00
SNU-4139S4-TO	TCF7L2	chr10	c.1496T>A	p.L499Q	0.27
SNU-4139S4-TO	KRAS	chr12	c.35G>T	p.G12V	0.25
SNU-4146_CT	MSH3	chr5	c.82T>G	p.F28V	0.27
SNU-4146_CT	APC	chr5	c.637C>T	p.R213*	0.51
SNU-4146_CT	APC	chr5	c.1312+2T>G	p.X438_splice	0.43
SNU-4146_CT	TP53	chr17	c.817C>T	p.R273C	0.73
SNU-4146S1-TO	APC	chr5	c.637C>T	p.R213*	0.45
SNU-4146S1-TO	APC	chr5	c.1312+2T>G	p.X438_splice	0.54
SNU-4146S1-TO	TP53	chr17	c.817C>T	p.R273C	1.00
SNU-4146S1T	APC	chr5	c.637C>T	p.R213*	0.37
SNU-4146S1T	APC	chr5	c.1312+2T>G	p.X438_splice	0.62
SNU-4146S1T	TP53	chr17	c.817C>T	p.R273C	1.00
SNU-4146S2-TO	APC	chr5	c.637C>T	p.R213*	0.71
SNU-4146S2-TO	APC	chr5	c.1312+2T>G	p.X438_splice	0.37
SNU-4146S2-TO	TP53	chr17	c.817C>T	p.R273C	1.00
SNU-4146S2	APC	chr5	c.637C>T	p.R213*	0.45
SNU-4146S2	APC	chr5	c.1312+2T>G	p.X438_splice	0.36
SNU-4146S2	TP53	chr17	c.817C>T	p.R273C	1.00
SNU-4146S3-TO	APC	chr5	c.637C>T	p.R213*	0.82
SNU-4146S3-TO	APC	chr5	c.1312+2T>G	p.X438_splice	0.26
SNU-4146S3-TO	TP53	chr17	c.817C>T	p.R273C	1.00
SNU-4146S3	APC	chr5	c.637C>T	p.R213*	0.28
SNU-4146S3	APC	chr5	c.1312+2T>G	p.X438_splice	0.69
SNU-4146S3	TP53	chr17	c.817C>T	p.R273C	1.00
SNU-4146S4-TO	APC	chr5	c.637C>T	p.R213*	0.53
SNU-4146S4-TO	APC	chr5	c.1312+2T>G	p.X438_splice	0.39
SNU-4146S4-TO	TP53	chr17	c.817C>T	p.R273C	1.00
SNU-4146S4T	APC	chr5	c.637C>T	p.R213*	0.42
SNU-4146S4T	APC	chr5	c.1312+2T>G	p.X438_splice	0.53
SNU-4146S4T	TP53	chr17	c.817C>T	p.R273C	1.00
SNU-4146S4	APC	chr5	c.637C>T	p.R213*	0.46
SNU-4146S4	APC	chr5	c.1312+2T>G	p.X438_splice	0.40
SNU-4146S4	TP53	chr17	c.817C>T	p.R273C	1.00

Continued

SampleCode	Hugo_Symbol	Chromosome	HGVSc	HGVSp_Short	VAF
SNU-4351_CT	APC	chr5	c.637C>T	p.R213*	0.33
SNU-4351_CT	APC	chr5	c.3925G>T	p.E1309*	0.58
SNU-4351_CT	KRAS	chr12	c.35G>T	p.G12V	0.53
SNU-4351_CT	AXIN2	chr17	c.289A>T	p.T97S	0.26
SNU-4351S1-TO	PIK3CA	chr3	c.1624G>A	p.E542K	0.84
SNU-4351S1-TO	APC	chr5	c.637C>T	p.R213*	0.51
SNU-4351S1-TO	APC	chr5	c.3925G>T	p.E1309*	0.47
SNU-4351S1-TO	KRAS	chr12	c.35G>T	p.G12V	0.68
SNU-4351S1-TO	AXIN2	chr17	c.289A>T	p.T97S	0.36
SNU-4351S2-TO	PIK3CA	chr3	c.1624G>A	p.E542K	0.73
SNU-4351S2-TO	APC	chr5	c.637C>T	p.R213*	0.26
SNU-4351S2-TO	APC	chr5	c.3925G>T	p.E1309*	0.50
SNU-4351S2-TO	KRAS	chr12	c.35G>T	p.G12V	0.62
SNU-4351S2-TO	AXIN2	chr17	c.289A>T	p.T97S	0.33
SNU-4351S3-TO	PIK3CA	chr3	c.1624G>A	p.E542K	0.63
SNU-4351S3-TO	APC	chr5	c.637C>T	p.R213*	0.38
SNU-4351S3-TO	APC	chr5	c.3925G>T	p.E1309*	0.60
SNU-4351S3-TO	KRAS	chr12	c.35G>T	p.G12V	0.69
SNU-4351S3-TO	AXIN2	chr17	c.289A>T	p.T97S	0.37
SNU-4351S4-TO	PIK3CA	chr3	c.1624G>A	p.E542K	0.82
SNU-4351S4-TO	APC	chr5	c.637C>T	p.R213*	0.31
SNU-4351S4-TO	APC	chr5	c.3925G>T	p.E1309*	0.55
SNU-4351S4-TO	KRAS	chr12	c.35G>T	p.G12V	0.63
SNU-4351S4-TO	AXIN2	chr17	c.289A>T	p.T97S	0.33
SNU-4374_CT	APC	chr5	c.4132C>T	p.Q1378*	0.66
SNU-4374_CT	TP53	chr17	c.406C>G	p.Q136E	0.46
SNU-4374S1-TO	APC	chr5	c.4132C>T	p.Q1378*	1.00
SNU-4374S1-TO	TP53	chr17	c.406C>G	p.Q136E	1.00
SNU-4374S2-TO	PIK3CA	chr3	c.1624G>A	p.E542K	0.73
SNU-4374S2-TO	MSH3	chr5	c.235A>G	p.I79V	0.60
SNU-4374S2-TO	MSH3	chr5	c.3133G>A	p.A1045T	0.49
SNU-4374S2-TO	APC	chr5	c.637C>T	p.R213*	0.26
SNU-4374S2-TO	APC	chr5	c.3925G>T	p.E1309*	0.50
SNU-4374S2-TO	KRAS	chr12	c.35G>T	p.G12V	0.62
SNU-4374S2-TO	TP53	chr17	c.91G>A	p.V31I	0.37
SNU-4374S2-TO	AXIN2	chr17	c.289A>T	p.T97S	0.33
SNU-4374S3-TO	APC	chr5	c.4132C>T	p.Q1378*	1.00
SNU-4374S3-TO	APC	chr5	c.8436C>A	p.N2812K	0.10
SNU-4374S3-TO	TP53	chr17	c.406C>G	p.Q136E	1.00
SNU-4374S4-TO	APC	chr5	c.4132C>T	p.Q1378*	1.00
SNU-4374S4-TO	TP53	chr17	c.406C>G	p.Q136E	1.00

Continued

SampleCode	Hugo_Symbol	Chromosome	HGVSc	HGVSp_Short	VAF
SNU-4376AS1-TO	ARID1A	chr1	c.1650del	p.Y551Tfs*68	0.31
SNU-4376AS1-TO	ARID1A	chr1	c.2017C>T	p.Q673*	0.36
SNU-4376AS1-TO	PIK3CA	chr3	c.3140A>G	p.H1047R	0.47
SNU-4376AS1-TO	APC	chr5	c.2098del	p.D700Tfs*18	0.51
SNU-4376AS1-TO	APC	chr5	c.4393_4394dup	p.S1465Rfs*9	0.43
SNU-4376AS1-TO	TCF7L2	chr10	c.1130G>A	p.S377N	0.56
SNU-4376AS1-TO	TCF7L2	chr10	c.1403del	p.K468Sfs*23	0.36
SNU-4376AS1-TO	KRAS	chr12	c.38G>A	p.G13D	0.60
SNU-4376AS1-TO	POLE	chr12	c.556G>A	p.A186T	0.49
SNU-4376AS1-TO	TP53	chr17	c.842A>G	p.D281G	0.58
SNU-4376AS1-TO	TP53	chr17	c.373A>G	p.T125A	0.77
SNU-4376AS1-TO	AXIN2	chr17	c.2011del	p.R671Afs*18	0.47
SNU-4376AS1T	ARID1A	chr1	c.1650del	p.Y551Tfs*68	0.63
SNU-4376AS1T	ARID1A	chr1	c.2017C>T	p.Q673*	0.52
SNU-4376AS1T	PIK3CA	chr3	c.3140A>G	p.H1047R	0.37
SNU-4376AS1T	APC	chr5	c.2098del	p.D700Tfs*18	0.36
SNU-4376AS1T	APC	chr5	c.4393_4394dup	p.S1465Rfs*9	0.45
SNU-4376AS1T	TCF7L2	chr10	c.1130G>A	p.S377N	0.76
SNU-4376AS1T	TCF7L2	chr10	c.1403del	p.K468Sfs*23	0.41
SNU-4376AS1T	KRAS	chr12	c.38G>A	p.G13D	0.65
SNU-4376AS1T	POLE	chr12	c.556G>A	p.A186T	0.61
SNU-4376AS1T	TP53	chr17	c.842A>G	p.D281G	0.54
SNU-4376AS1T	TP53	chr17	c.373A>G	p.T125A	0.37
SNU-4376AS1T	AXIN2	chr17	c.2011del	p.R671Afs*18	0.45
SNU-4376AS1	ARID1A	chr1	c.1650del	p.Y551Tfs*68	0.57
SNU-4376AS1	ARID1A	chr1	c.2017C>T	p.Q673*	0.53
SNU-4376AS1	PIK3CA	chr3	c.3140A>G	p.H1047R	0.54
SNU-4376AS1	APC	chr5	c.2098del	p.D700Tfs*18	0.54
SNU-4376AS1	APC	chr5	c.4393_4394dup	p.S1465Rfs*9	0.56
SNU-4376AS1	TCF7L2	chr10	c.1130G>A	p.S377N	0.44
SNU-4376AS1	TCF7L2	chr10	c.1403del	p.K468Sfs*23	0.54
SNU-4376AS1	KRAS	chr12	c.38G>A	p.G13D	0.38
SNU-4376AS1	POLE	chr12	c.556G>A	p.A186T	0.64
SNU-4376AS1	TP53	chr17	c.842A>G	p.D281G	0.53
SNU-4376AS1	TP53	chr17	c.373A>G	p.T125A	0.50
SNU-4376AS1	AXIN2	chr17	c.2011del	p.R671Afs*18	0.47
SNU-4376AS2	ARID1A	chr1	c.1650del	p.Y551Tfs*68	0.53
SNU-4376AS2	ARID1A	chr1	c.2017C>T	p.Q673*	0.48
SNU-4376AS2	PIK3CA	chr3	c.3140A>G	p.H1047R	0.50
SNU-4376AS2	APC	chr5	c.2098del	p.D700Tfs*18	0.55
SNU-4376AS2	APC	chr5	c.4393_4394dup	p.S1465Rfs*9	0.60
SNU-4376AS2	TCF7L2	chr10	c.1130G>A	p.S377N	0.57
SNU-4376AS2	TCF7L2	chr10	c.1403del	p.K468Sfs*23	0.49
SNU-4376AS2	KRAS	chr12	c.38G>A	p.G13D	0.61
SNU-4376AS2	ARID2	chr12	c.1739T>C	p.V580A	0.79
SNU-4376AS2	POLE	chr12	c.556G>A	p.A186T	0.47
SNU-4376AS2	TP53	chr17	c.842A>G	p.D281G	0.48
SNU-4376AS2	TP53	chr17	c.373A>G	p.T125A	0.61
SNU-4376AS2	AXIN2	chr17	c.2011del	p.R671Afs*18	0.44

Continued

SampleCode	Hugo_Symbol	Chromosome	HGVSc	HGVSp_Short	VAF
SNU-4376AS3-TO	ARID1A	chr1	c.1650del	p.Y551Tfs*68	0.51
SNU-4376AS3-TO	ARID1A	chr1	c.2017C>T	p.Q673*	0.49
SNU-4376AS3-TO	PIK3CA	chr3	c.3140A>G	p.H1047R	0.50
SNU-4376AS3-TO	APC	chr5	c.2098del	p.D700Tfs*18	0.51
SNU-4376AS3-TO	APC	chr5	c.4393_4394dup	p.S1465Rfs*9	0.52
SNU-4376AS3-TO	TCF7L2	chr10	c.1130G>A	p.S377N	0.56
SNU-4376AS3-TO	TCF7L2	chr10	c.1403del	p.K468Sfs*23	0.45
SNU-4376AS3-TO	KRAS	chr12	c.38G>A	p.G13D	0.64
SNU-4376AS3-TO	POLE	chr12	c.556G>A	p.A186T	0.44
SNU-4376AS3-TO	TP53	chr17	c.842A>G	p.D281G	0.61
SNU-4376AS3-TO	TP53	chr17	c.373A>G	p.T125A	0.44
SNU-4376AS3-TO	AXIN2	chr17	c.2011del	p.R671Afs*18	0.47
SNU-4376AS3T	ARID1A	chr1	c.1650del	p.Y551Tfs*68	0.53
SNU-4376AS3T	ARID1A	chr1	c.2017C>T	p.Q673*	0.46
SNU-4376AS3T	PIK3CA	chr3	c.3140A>G	p.H1047R	0.62
SNU-4376AS3T	APC	chr5	c.2098del	p.D700Tfs*18	0.41
SNU-4376AS3T	APC	chr5	c.4393_4394dup	p.S1465Rfs*9	0.41
SNU-4376AS3T	TCF7L2	chr10	c.1130G>A	p.S377N	0.45
SNU-4376AS3T	TCF7L2	chr10	c.1403del	p.K468Sfs*23	0.61
SNU-4376AS3T	KRAS	chr12	c.38G>A	p.G13D	0.56
SNU-4376AS3T	POLE	chr12	c.556G>A	p.A186T	0.49
SNU-4376AS3T	TP53	chr17	c.842A>G	p.D281G	0.50
SNU-4376AS3T	TP53	chr17	c.373A>G	p.T125A	0.46
SNU-4376AS3T	AXIN2	chr17	c.2011del	p.R671Afs*18	0.51
SNU-4376AS3	ARID1A	chr1	c.1650del	p.Y551Tfs*68	0.38
SNU-4376AS3	ARID1A	chr1	c.2017C>T	p.Q673*	0.53
SNU-4376AS3	PIK3CA	chr3	c.3140A>G	p.H1047R	0.58
SNU-4376AS3	APC	chr5	c.2098del	p.D700Tfs*18	0.51
SNU-4376AS3	APC	chr5	c.4393_4394dup	p.S1465Rfs*9	0.52
SNU-4376AS3	TCF7L2	chr10	c.1130G>A	p.S377N	0.50
SNU-4376AS3	TCF7L2	chr10	c.1403del	p.K468Sfs*23	0.52
SNU-4376AS3	KRAS	chr12	c.38G>A	p.G13D	0.40
SNU-4376AS3	POLE	chr12	c.556G>A	p.A186T	0.53
SNU-4376AS3	TP53	chr17	c.842A>G	p.D281G	0.47
SNU-4376AS3	TP53	chr17	c.373A>G	p.T125A	0.57
SNU-4376AS3	AXIN2	chr17	c.2011del	p.R671Afs*18	0.44
SNU-4376AS4-TO	ARID1A	chr1	c.1650del	p.Y551Tfs*68	0.55
SNU-4376AS4-TO	ARID1A	chr1	c.2017C>T	p.Q673*	0.55
SNU-4376AS4-TO	ARID1A	chr1	c.5371del	p.S1791Qfs*15	0.71
SNU-4376AS4-TO	PIK3CA	chr3	c.3140A>G	p.H1047R	0.50
SNU-4376AS4-TO	APC	chr5	c.2098del	p.D700Tfs*18	0.46
SNU-4376AS4-TO	APC	chr5	c.4393_4394dup	p.S1465Rfs*9	0.50
SNU-4376AS4-TO	TCF7L2	chr10	c.1130G>A	p.S377N	0.63
SNU-4376AS4-TO	TCF7L2	chr10	c.1403del	p.K468Sfs*23	0.48
SNU-4376AS4-TO	KRAS	chr12	c.38G>A	p.G13D	0.67
SNU-4376AS4-TO	POLE	chr12	c.556G>A	p.A186T	0.55
SNU-4376AS4-TO	TP53	chr17	c.842A>G	p.D281G	0.59
SNU-4376AS4-TO	TP53	chr17	c.373A>G	p.T125A	0.36
SNU-4376AS4-TO	AXIN2	chr17	c.2011del	p.R671Afs*18	0.48
SNU-4376AS4-TO	SMAD4	chr18	c.88G>A	p.G30R	0.73
SNU-4376AS4	ARID1A	chr1	c.1650del	p.Y551Tfs*68	0.45
SNU-4376AS4	ARID1A	chr1	c.2017C>T	p.Q673*	0.53
SNU-4376AS4	ARID1A	chr1	c.5371del	p.S1791Qfs*15	0.46
SNU-4376AS4	PIK3CA	chr3	c.3140A>G	p.H1047R	0.54
SNU-4376AS4	APC	chr5	c.2098del	p.D700Tfs*18	0.44
SNU-4376AS4	APC	chr5	c.4393_4394dup	p.S1465Rfs*9	0.56
SNU-4376AS4	TCF7L2	chr10	c.1130G>A	p.S377N	0.43
SNU-4376AS4	TCF7L2	chr10	c.1403del	p.K468Sfs*23	0.41
SNU-4376AS4	KRAS	chr12	c.38G>A	p.G13D	0.48
SNU-4376AS4	POLE	chr12	c.556G>A	p.A186T	0.53
SNU-4376AS4	TP53	chr17	c.842A>G	p.D281G	0.44
SNU-4376AS4	TP53	chr17	c.373A>G	p.T125A	0.72
SNU-4376AS4	AXIN2	chr17	c.2011del	p.R671Afs*18	0.54
SNU-4376AS4	SMAD4	chr18	c.88G>A	p.G30R	0.61

Continued

SampleCode	Hugo_Symbol	Chromosome	HGVSc	HGVSp_Short	VAF
SNU-4398_CT	ARID1A	chr1	c.827del	p.G276Efs*87	0.43
SNU-4398_CT	ARID1A	chr1	c.3910G>A	p.A1304T	0.35
SNU-4398_CT	ARID1A	chr1	c.4196A>G	p.Q1399R	0.39
SNU-4398_CT	PIK3CA	chr3	c.320A>G	p.N107S	0.43
SNU-4398_CT	FBXW7	chr4	c.1629_1630del	p.R543Sfs*7	0.28
SNU-4398_CT	MSH3	chr5	c.994G>A	p.A332T	0.36
SNU-4398_CT	APC	chr5	c.1742dup	p.E582Gfs*20	0.46
SNU-4398_CT	APC	chr5	c.4666dup	p.T1556Nfs*3	0.33
SNU-4398_CT	APC	chr5	c.7749del	p.A2584Qfs*9	0.47
SNU-4398_CT	POLE	chr12	c.122C>T	p.T41M	0.36
SNU-4398_CT	AXIN2	chr17	c.1994del	p.G665Afs*24	0.29
SNU-4398S1-TO	ARID1A	chr1	c.3910G>A	p.A1304T	0.20
SNU-4398S1-TO	ARID1A	chr1	c.4196A>G	p.Q1399R	0.43
SNU-4398S1-TO	PIK3CA	chr3	c.320A>G	p.N107S	0.48
SNU-4398S1-TO	FBXW7	chr4	c.1629_1630del	p.R543Sfs*7	0.45
SNU-4398S1-TO	MSH3	chr5	c.994G>A	p.A332T	0.45
SNU-4398S1-TO	APC	chr5	c.1742dup	p.E582Gfs*20	0.50
SNU-4398S1-TO	APC	chr5	c.4666dup	p.T1556Nfs*3	0.34
SNU-4398S1-TO	APC	chr5	c.7749del	p.A2584Qfs*9	0.55
SNU-4398S1-TO	POLE	chr12	c.122C>T	p.T41M	0.50
SNU-4398S1-TO	AXIN2	chr17	c.1994del	p.G665Afs*24	0.58
SNU-4398S1	ARID1A	chr1	c.1466C>T	p.P489L	0.34
SNU-4398S1	ARID1A	chr1	c.4196A>G	p.Q1399R	0.46
SNU-4398S1	FBXW7	chr4	c.1629_1630del	p.R543Sfs*7	0.46
SNU-4398S1	MSH3	chr5	c.994G>A	p.A332T	0.44
SNU-4398S1	APC	chr5	c.1742dup	p.E582Gfs*20	0.55
SNU-4398S1	APC	chr5	c.4666dup	p.T1556Nfs*3	0.53
SNU-4398S1	TCF7L2	chr10	c.1403del	p.K468Sfs*23	0.33
SNU-4398S1	POLE	chr12	c.122C>T	p.T41M	0.42
SNU-4398S1	AXIN2	chr17	c.1994del	p.G665Afs*24	0.52
SNU-4398S2-TO	ARID1A	chr1	c.4196A>G	p.Q1399R	0.48
SNU-4398S2-TO	FBXW7	chr4	c.1629_1630del	p.R543Sfs*7	0.45
SNU-4398S2-TO	MSH3	chr5	c.994G>A	p.A332T	0.49
SNU-4398S2-TO	APC	chr5	c.1742dup	p.E582Gfs*20	0.60
SNU-4398S2-TO	APC	chr5	c.4666dup	p.T1556Nfs*3	0.49
SNU-4398S2-TO	TCF7L2	chr10	c.14A>G	p.N5S	0.21
SNU-4398S2-TO	ERBB3	chr12	c.2578C>T	p.P860S	0.26
SNU-4398S2-TO	POLE	chr12	c.122C>T	p.T41M	0.53
SNU-4398S2-TO	AXIN2	chr17	c.1994del	p.G665Afs*24	0.51
SNU-4398S2-TO	SMAD2	chr18	c.341G>A	p.R114H	0.15

Continued

SampleCode	Hugo_Symbol	Chromosome	HGVSc	HGVSp_Short	VAF
SNU-4398S2	ARID1A	chr1	c.4196A>G	p.Q1399R	0.51
SNU-4398S2	PIK3CA	chr3	c.3131A>G	p.N1044S	0.44
SNU-4398S2	FBXW7	chr4	c.1629_1630del	p.R543Sfs*7	0.42
SNU-4398S2	MSH3	chr5	c.994G>A	p.A332T	0.51
SNU-4398S2	MSH3	chr5	c.1148del	p.K383Rfs*32	0.49
SNU-4398S2	APC	chr5	c.1742dup	p.E582Gfs*20	0.43
SNU-4398S2	APC	chr5	c.4666dup	p.T1556Nfs*3	0.47
SNU-4398S2	POLE	chr12	c.122C>T	p.T41M	0.56
SNU-4398S2	AXIN2	chr17	c.1994del	p.G665Afs*24	0.44
SNU-4398S3-TO	ARID1A	chr1	c.4196A>G	p.Q1399R	0.51
SNU-4398S3-TO	PIK3CA	chr3	c.320A>G	p.N107S	0.48
SNU-4398S3-TO	FBXW7	chr4	c.1629_1630del	p.R543Sfs*7	0.47
SNU-4398S3-TO	MSH3	chr5	c.994G>A	p.A332T	0.43
SNU-4398S3-TO	APC	chr5	c.1742dup	p.E582Gfs*20	0.53
SNU-4398S3-TO	APC	chr5	c.4666dup	p.T1556Nfs*3	0.38
SNU-4398S3-TO	APC	chr5	c.7749del	p.A2584Qfs*9	0.47
SNU-4398S3-TO	POLE	chr12	c.122C>T	p.T41M	0.45
SNU-4398S3-TO	AXIN2	chr17	c.1994del	p.G665Afs*24	0.54
SNU-4398S4-TO	ARID1A	chr1	c.4196A>G	p.Q1399R	0.49
SNU-4398S4-TO	CTNNB1	chr3	c.830G>A	p.G277D	0.16
SNU-4398S4-TO	PIK3CA	chr3	c.2816A>G	p.D939G	0.36
SNU-4398S4-TO	FBXW7	chr4	c.1629_1630del	p.R543Sfs*7	0.54
SNU-4398S4-TO	MSH3	chr5	c.994G>A	p.A332T	0.38
SNU-4398S4-TO	APC	chr5	c.1742dup	p.E582Gfs*20	0.50
SNU-4398S4-TO	APC	chr5	c.4666dup	p.T1556Nfs*3	0.58
SNU-4398S4-TO	POLE	chr12	c.122C>T	p.T41M	0.48
SNU-4398S4-TO	AXIN2	chr17	c.1994del	p.G665Afs*24	0.47
SNU-4398S4	ARID1A	chr1	c.4196A>G	p.Q1399R	0.50
SNU-4398S4	PIK3CA	chr3	c.2816A>G	p.D939G	0.49
SNU-4398S4	FBXW7	chr4	c.1629_1630del	p.R543Sfs*7	0.59
SNU-4398S4	MSH3	chr5	c.994G>A	p.A332T	0.62
SNU-4398S4	APC	chr5	c.1742dup	p.E582Gfs*20	0.58
SNU-4398S4	APC	chr5	c.4666dup	p.T1556Nfs*3	0.56
SNU-4398S4	TCF7L2	chr10	c.1403del	p.K468Sfs*23	0.21
SNU-4398S4	POLE	chr12	c.122C>T	p.T41M	0.45
SNU-4398S4	TP53	chr17	c.481G>A	p.A161T	0.29
SNU-4398S4	AXIN2	chr17	c.1994del	p.G665Afs*24	0.49

Continued

SampleCode	Hugo_Symbol	Chromosome	HGVSc	HGVSp_Short	VAF
SNU-4631A_CT	ARID1A	chr1	c.5893A>C	p.T1965P	0.23
SNU-4631A_CT	PIK3CA	chr3	c.1633G>A	p.E545K	0.37
SNU-4631A_CT	FBXW7	chr4	c.1744T>C	p.S582P	0.60
SNU-4631A_CT	TP53	chr17	c.517G>T	p.V173L	0.45
SNU-4631AS1-TO	ARID1A	chr1	c.5893A>C	p.T1965P	0.50
SNU-4631AS1-TO	PIK3CA	chr3	c.1633G>A	p.E545K	0.44
SNU-4631AS1-TO	FBXW7	chr4	c.1744T>C	p.S582P	1.00
SNU-4631AS1-TO	TP53	chr17	c.517G>T	p.V173L	1.00
SNU-4631AS1	ARID1A	chr1	c.5893A>C	p.T1965P	0.47
SNU-4631AS1	PIK3CA	chr3	c.1633G>A	p.E545K	0.56
SNU-4631AS1	FBXW7	chr4	c.1744T>C	p.S582P	1.00
SNU-4631AS1	TP53	chr17	c.517G>T	p.V173L	1.00
SNU-4631AS2-TO	ARID1A	chr1	c.5893A>C	p.T1965P	0.50
SNU-4631AS2-TO	PIK3CA	chr3	c.1633G>A	p.E545K	0.75
SNU-4631AS2-TO	FBXW7	chr4	c.1744T>C	p.S582P	1.00
SNU-4631AS2-TO	TP53	chr17	c.517G>T	p.V173L	1.00
SNU-4631AS3-TO	ARID1A	chr1	c.5893A>C	p.T1965P	0.50
SNU-4631AS3-TO	PIK3CA	chr3	c.1633G>A	p.E545K	0.78
SNU-4631AS3-TO	FBXW7	chr4	c.1744T>C	p.S582P	1.00
SNU-4631AS3-TO	TP53	chr17	c.517G>T	p.V173L	1.00
SNU-4631AS4-TO	ARID1A	chr1	c.5893A>C	p.T1965P	0.48
SNU-4631AS4-TO	PIK3CA	chr3	c.1633G>A	p.E545K	0.59
SNU-4631AS4-TO	FBXW7	chr4	c.1744T>C	p.S582P	0.96
SNU-4631AS4-TO	TP53	chr17	c.517G>T	p.V173L	1.00
SNU-4631AS4	ARID1A	chr1	c.5893A>C	p.T1965P	0.45
SNU-4631AS4	PIK3CA	chr3	c.1633G>A	p.E545K	0.50
SNU-4631AS4	FBXW7	chr4	c.1744T>C	p.S582P	1.00
SNU-4631AS4	TP53	chr17	c.517G>T	p.V173L	1.00
SNU-4646_CT	PIK3CA	chr3	c.3141T>A	p.H1047Q	0.776786
SNU-4646_CT	APC	chr5	c.847C>T	p.R283*	0.735849
SNU-4646_CT	KRAS	chr12	c.35G>A	p.G12D	0.854545
SNU-4646_CT	TP53	chr17	c.524G>A	p.R175H	0.753247
SNU-4646S1-TO	PIK3CA	chr3	c.3141T>A	p.H1047Q	0.986301
SNU-4646S1-TO	APC	chr5	c.847C>T	p.R283*	0.925
SNU-4646S1-TO	KRAS	chr12	c.35G>A	p.G12D	0.454545
SNU-4646S1-TO	TP53	chr17	c.524G>A	p.R175H	0.963855
SNU-4646S1T	PIK3CA	chr3	c.3141T>A	p.H1047Q	0.896552
SNU-4646S1T	APC	chr5	c.847C>T	p.R283*	0.821429
SNU-4646S1T	KRAS	chr12	c.35G>A	p.G12D	0.681818
SNU-4646S1T	TP53	chr17	c.524G>A	p.R175H	0.987805
SNU-4646S2-TO	PIK3CA	chr3	c.3141T>A	p.H1047Q	0.961039
SNU-4646S2-TO	FBXW7	chr4	c.176T>A	p.V59D	0.846154
SNU-4646S2-TO	APC	chr5	c.847C>T	p.R283*	0.964286
SNU-4646S2-TO	KRAS	chr12	c.35G>A	p.G12D	0.375
SNU-4646S2-TO	TP53	chr17	c.524G>A	p.R175H	0.958904
SNU-4646S2T	PIK3CA	chr3	c.3141T>A	p.H1047Q	0.962963
SNU-4646S2T	APC	chr5	c.847C>T	p.R283*	1
SNU-4646S2T	KRAS	chr12	c.35G>A	p.G12D	0.411765
SNU-4646S2T	TP53	chr17	c.524G>A	p.R175H	0.971831
SNU-4646S3-TO	PIK3CA	chr3	c.3141T>A	p.H1047Q	0.065217
SNU-4646S3-TO	APC	chr5	c.847C>T	p.R283*	1
SNU-4646S3-TO	KRAS	chr12	c.35G>A	p.G12D	0.4
SNU-4646S3-TO	TP53	chr17	c.524G>A	p.R175H	0.916667
SNU-4646S3T	PIK3CA	chr3	c.3141T>A	p.H1047Q	0.846154
SNU-4646S3T	APC	chr5	c.847C>T	p.R283*	1
SNU-4646S3T	KRAS	chr12	c.35G>A	p.G12D	0.277778
SNU-4646S3T	TP53	chr17	c.524G>A	p.R175H	0.970588

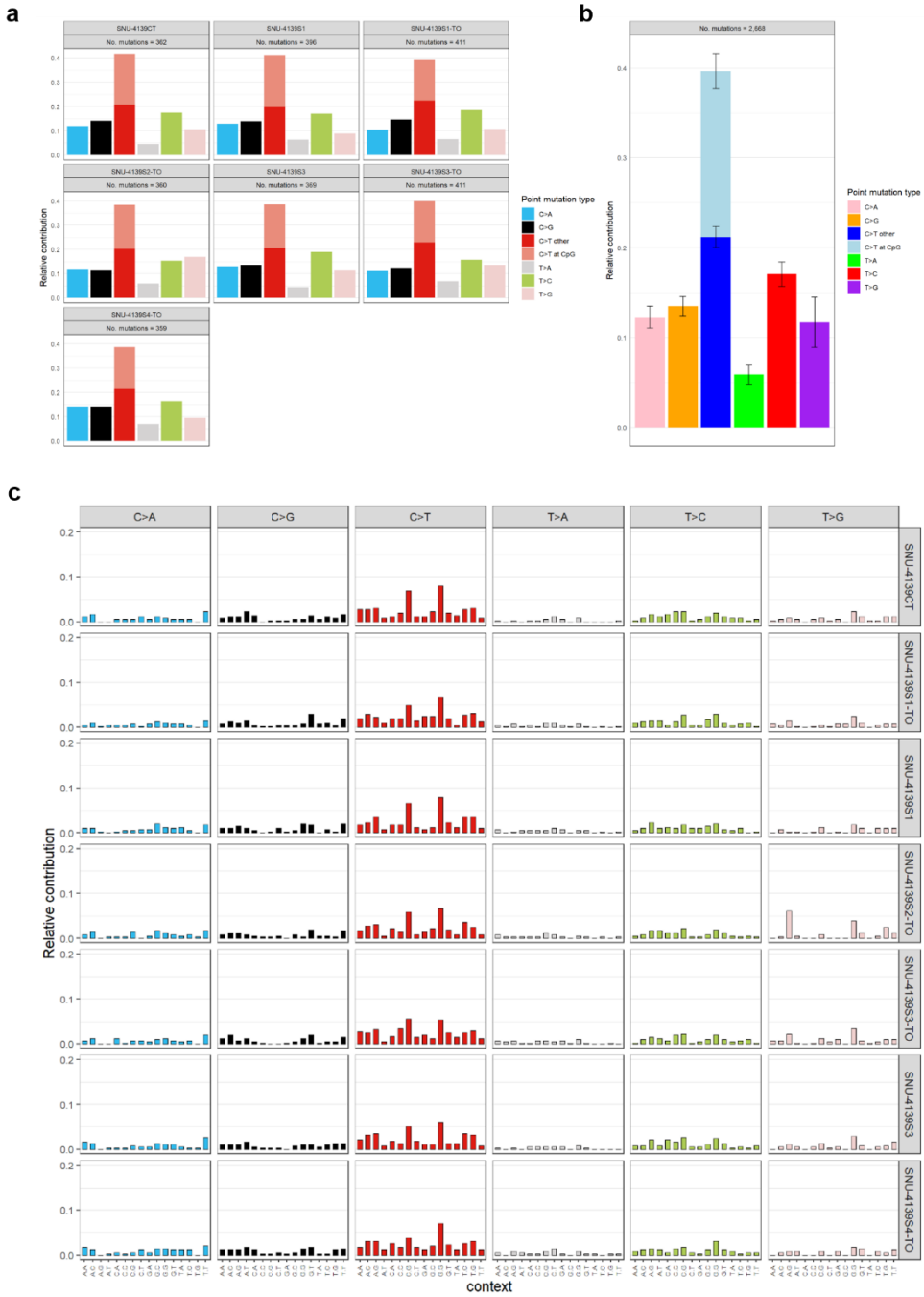
Continued

SampleCode	Hugo_Symbol	Chromosome	HGVSc	HGVSp_Short	VAF
SNU-4713_CT	FBXW7	chr4	c.1745C>T	p.S582L	0.644444
SNU-4713_CT	FBXW7	chr4	c.1685C>A	p.S562*	0.538462
SNU-4713_CT	APC	chr5	c.3334dup	p.T1112Nfs*7	0.822222
SNU-4713_CT	TP53	chr17	c.349_355dup	p.A119Gfs*32	0.878788
SNU-4713S1-TO	FBXW7	chr4	c.1745C>T	p.S582L	0.45
SNU-4713S1-TO	FBXW7	chr4	c.1685C>A	p.S562*	0.62
SNU-4713S1-TO	APC	chr5	c.3334dup	p.T1112Nfs*7	1.00
SNU-4713S1-TO	TP53	chr17	c.349_355dup	p.A119Gfs*32	1.00
SNU-4713S1T	FBXW7	chr4	c.1745C>T	p.S582L	0.68
SNU-4713S1T	FBXW7	chr4	c.1685C>A	p.S562*	0.28
SNU-4713S1T	APC	chr5	c.3334dup	p.T1112Nfs*7	1.00
SNU-4713S1T	TP53	chr17	c.349_355dup	p.A119Gfs*32	1.00
SNU-4713S1	FBXW7	chr4	c.1745C>T	p.S582L	0.66
SNU-4713S1	FBXW7	chr4	c.1685C>A	p.S562*	0.35
SNU-4713S1	APC	chr5	c.3334dup	p.T1112Nfs*7	0.96
SNU-4713S1	TP53	chr17	c.349_355dup	p.A119Gfs*32	0.96
SNU-4713S2-TO	FBXW7	chr4	c.1745C>T	p.S582L	0.98
SNU-4713S2-TO	APC	chr5	c.3334dup	p.T1112Nfs*7	1.00
SNU-4713S2-TO	TP53	chr17	c.349_355dup	p.A119Gfs*32	1.00
SNU-4713S2-TO	AXIN2	chr17	c.423C>A	p.N141K	0.71
SNU-4713S3-TO	FBXW7	chr4	c.1745C>T	p.S582L	0.23
SNU-4713S3-TO	FBXW7	chr4	c.1685C>A	p.S562*	0.71
SNU-4713S3-TO	APC	chr5	c.3334dup	p.T1112Nfs*7	1.00
SNU-4713S3-TO	TP53	chr17	c.349_355dup	p.A119Gfs*32	1.00
SNU-4713S3-TO	AXIN2	chr17	c.423C>A	p.N141K	0.59
SNU-4713S3	FBXW7	chr4	c.1745C>T	p.S582L	0.51
SNU-4713S3	FBXW7	chr4	c.1685C>A	p.S562*	0.49
SNU-4713S3	APC	chr5	c.3334dup	p.T1112Nfs*7	1.00
SNU-4713S3	TP53	chr17	c.349_355dup	p.A119Gfs*32	1.00
SNU-4713S3	AXIN2	chr17	c.423C>A	p.N141K	0.42
SNU-4796_CT	KRAS	chr12	c.40G>A	p.V14I	0.45
SNU-4796_CT	TP53	chr17	c.524G>A	p.R175H	0.45
SNU-4796S1-TO	KRAS	chr12	c.40G>A	p.V14I	0.61
SNU-4796S1-TO	TP53	chr17	c.524G>A	p.R175H	0.99
SNU-4796S2-TO	KRAS	chr12	c.40G>A	p.V14I	0.59
SNU-4796S2-TO	TP53	chr17	c.524G>A	p.R175H	1.00
SNU-4796S3-TO	KRAS	chr12	c.40G>A	p.V14I	0.54
SNU-4796S3-TO	TP53	chr17	c.524G>A	p.R175H	1.00
SNU-4796S4-TO	KRAS	chr12	c.40G>A	p.V14I	0.57
SNU-4796S4-TO	TP53	chr17	c.524G>A	p.R175H	1.00

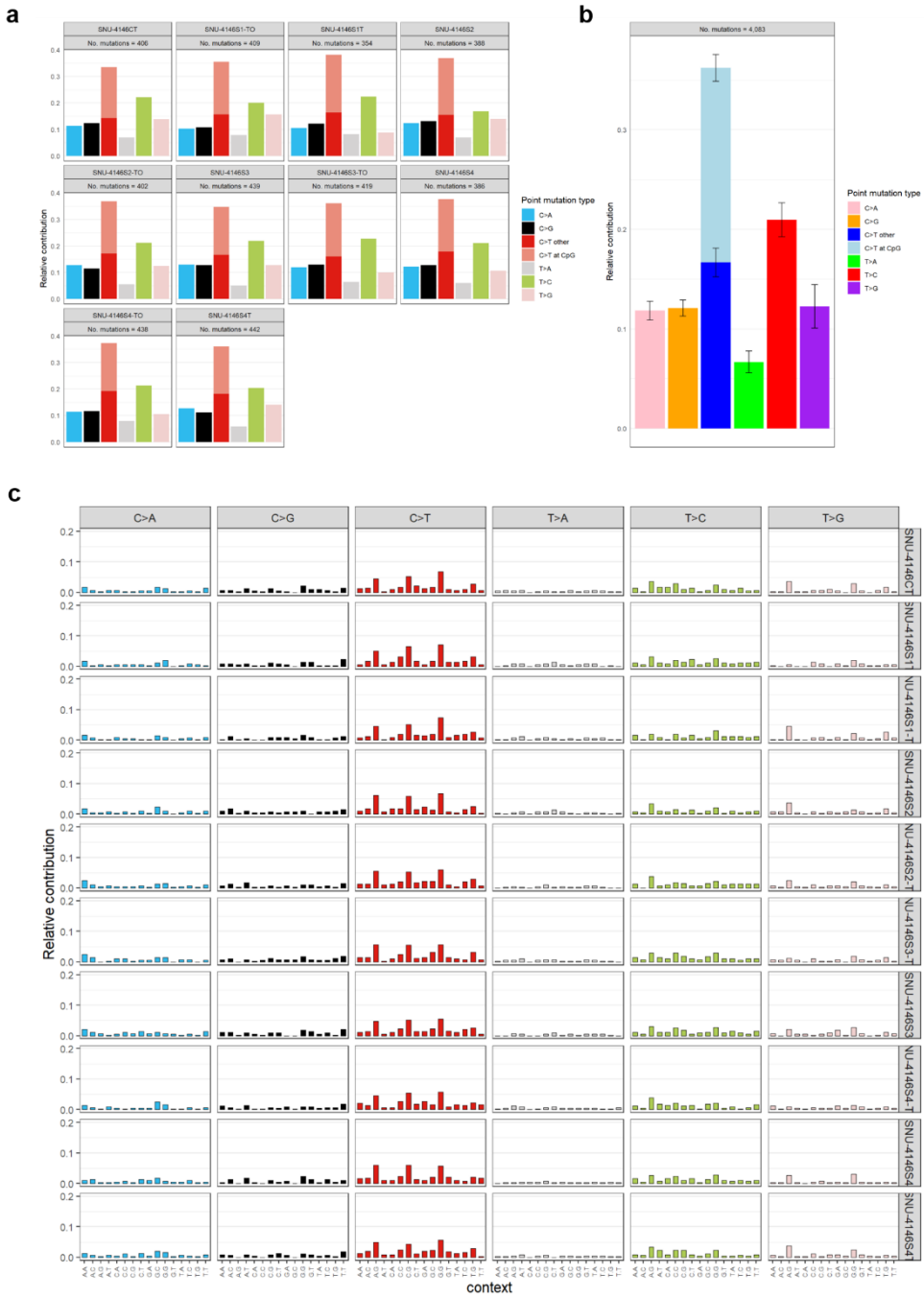
Continued

SampleCode	Hugo_Symbol	Chromosome	HGVSc	HGVSp_Short	VAF
SNU-4813_CT	FBXW7	chr4	c.1154C>T	p.T385I	0.53
SNU-4813_CT	APC	chr5	c.2735T>A	p.L912*	0.39
SNU-4813_CT	KRAS	chr12	c.76A>T	p.N26Y	0.18
SNU-4813_CT	TP53	chr17	c.389T>A	p.L130H	0.34
SNU-4813S1-TO	FBXW7	chr4	c.1154C>T	p.T385I	1.00
SNU-4813S1-TO	APC	chr5	c.2735T>A	p.L912*	0.98
SNU-4813S1-TO	KRAS	chr12	c.76A>T	p.N26Y	0.54
SNU-4813S1-TO	TP53	chr17	c.389T>A	p.L130H	1.00
SNU-4813S2-TO	FBXW7	chr4	c.1154C>T	p.T385I	1.00
SNU-4813S2-TO	APC	chr5	c.2735T>A	p.L912*	1.00
SNU-4813S2-TO	KRAS	chr12	c.76A>T	p.N26Y	0.60
SNU-4813S2-TO	TP53	chr17	c.389T>A	p.L130H	0.99
SNU-4813S3-TO	FBXW7	chr4	c.1154C>T	p.T385I	1.00
SNU-4813S3-TO	APC	chr5	c.2735T>A	p.L912*	0.97
SNU-4813S3-TO	KRAS	chr12	c.76A>T	p.N26Y	0.35
SNU-4813S3-TO	TP53	chr17	c.389T>A	p.L130H	1.00
SNU-4849_CT	ARID1A	chr1	c.4187_4188del	p.G1396Afs*48	0.49
SNU-4849_CT	APC	chr5	c.4132C>T	p.Q1378*	0.52
SNU-4849_CT	TP53	chr17	c.713G>A	p.C238Y	0.57
SNU-4849_CT	SMAD4	chr18	c.1217C>T	p.A406V	0.65
SNU-4849S1-TO	ARID1A	chr1	c.4187_4188del	p.G1396Afs*48	0.56
SNU-4849S1-TO	APC	chr5	c.4132C>T	p.Q1378*	1.00
SNU-4849S1-TO	TP53	chr17	c.713G>A	p.C238Y	0.94
SNU-4849S1-TO	SMAD4	chr18	c.1217C>T	p.A406V	1.00
SNU-4849S2-TO	ARID1A	chr1	c.4187_4188del	p.G1396Afs*48	0.52
SNU-4849S2-TO	APC	chr5	c.4132C>T	p.Q1378*	1.00
SNU-4849S2-TO	TP53	chr17	c.713G>A	p.C238Y	1.00
SNU-4849S2-TO	SMAD4	chr18	c.1217C>T	p.A406V	1.00
SNU-4849S3-TO	ARID1A	chr1	c.4187_4188del	p.G1396Afs*48	0.51
SNU-4849S3-TO	APC	chr5	c.4132C>T	p.Q1378*	0.94
SNU-4849S3-TO	TP53	chr17	c.713G>A	p.C238Y	0.98
SNU-4849S3-TO	SMAD4	chr18	c.1217C>T	p.A406V	1.00

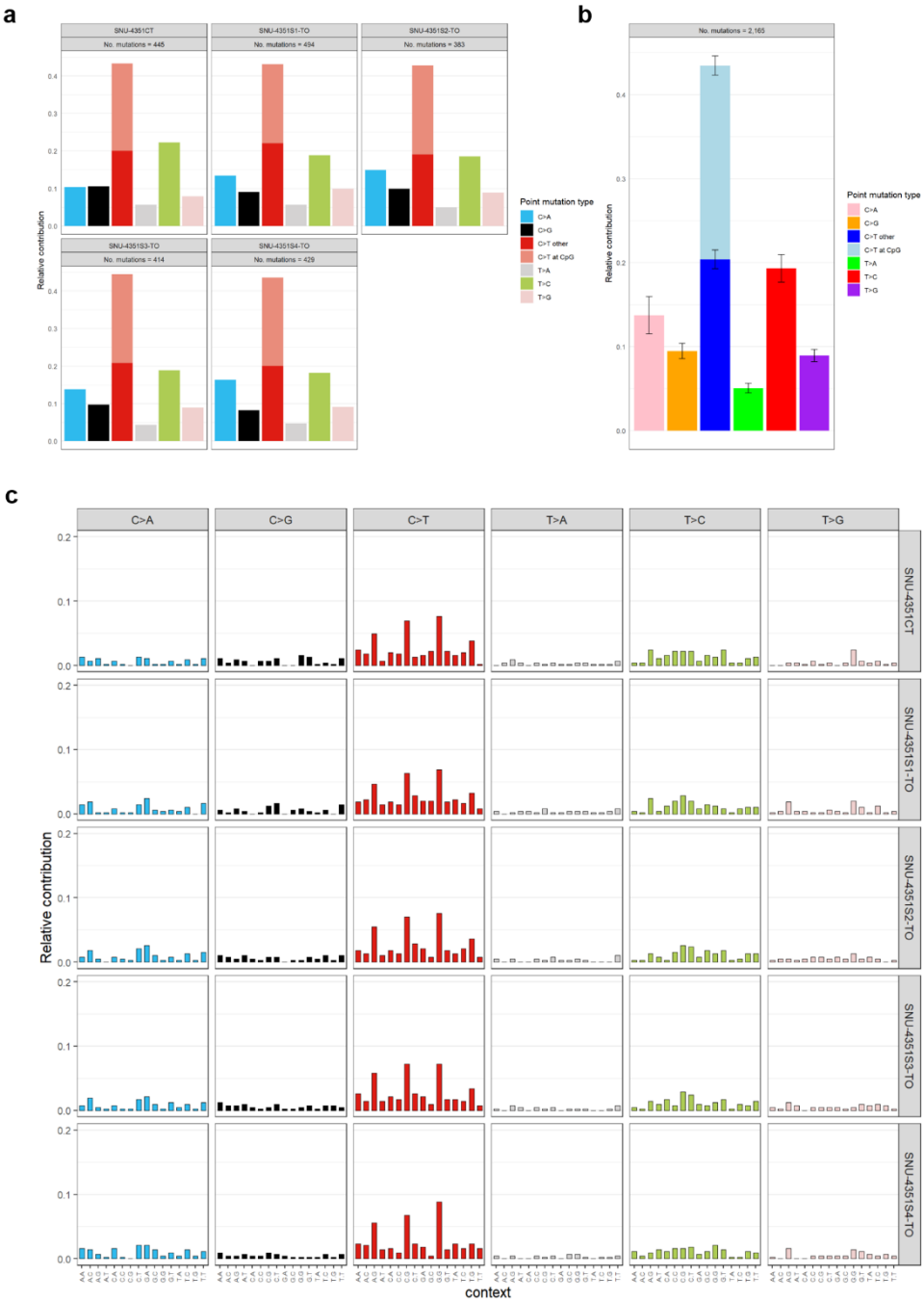
A. SNU-4139-TO


a
SNU-4139CT
No. mutations = 362
SNU-4139S1
No. mutations = 396
SNU-4139S1-TO
No. mutations = 411
SNU-4139S2-TO
No. mutations = 360
SNU-4139S3
No. mutations = 369
SNU-4139S3-TO
No. mutations = 411
SNU-4139S4-TO
No. mutations = 359
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
b
No. mutations = 2,668
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
c
C>A
C>G
C>T
T>A
T>C
T>G
SNU-4139CT
SNU-4139S1-TO
SNU-4139S1
SNU-4139S2-TO
SNU-4139S3-TO
SNU-4139S3
SNU-4139S4-TO
Relative contribution
context

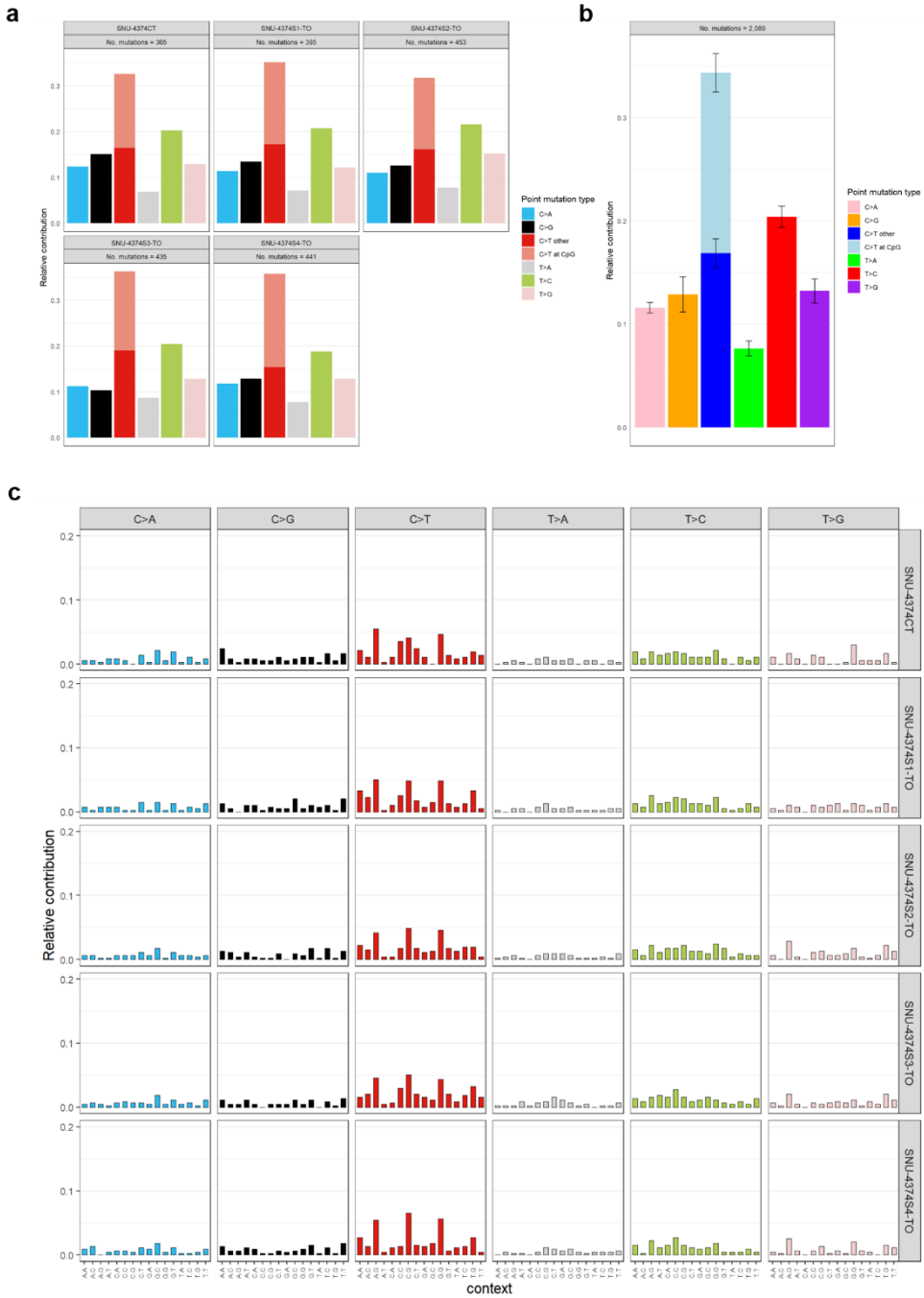
B. SNU-4146-TO


a
SNU-4146CT
No. mutations = 406
SNU-4146S1-TO
No. mutations = 409
SNU-4146S1T
No. mutations = 354
SNU-4146S2
No. mutations = 388
SNU-4146S2-TO
No. mutations = 402
SNU-4146S3
No. mutations = 439
SNU-4146S3-TO
No. mutations = 419
SNU-4146S4
No. mutations = 386
SNU-4146S4-TO
No. mutations = 438
SNU-4146S4T
No. mutations = 442
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
b
No. mutations = 4,083
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
c
C>A
C>G
C>T
T>A
T>C
T>G
SNU-4146CT
SNU-4146S1
SNU-4146S1-T
SNU-4146S2
SNU-4146S2-T
SNU-4146S3-T
SNU-4146S3
SNU-4146S4-T
SNU-4146S4
SNU-4146S4T
Relative contribution
context

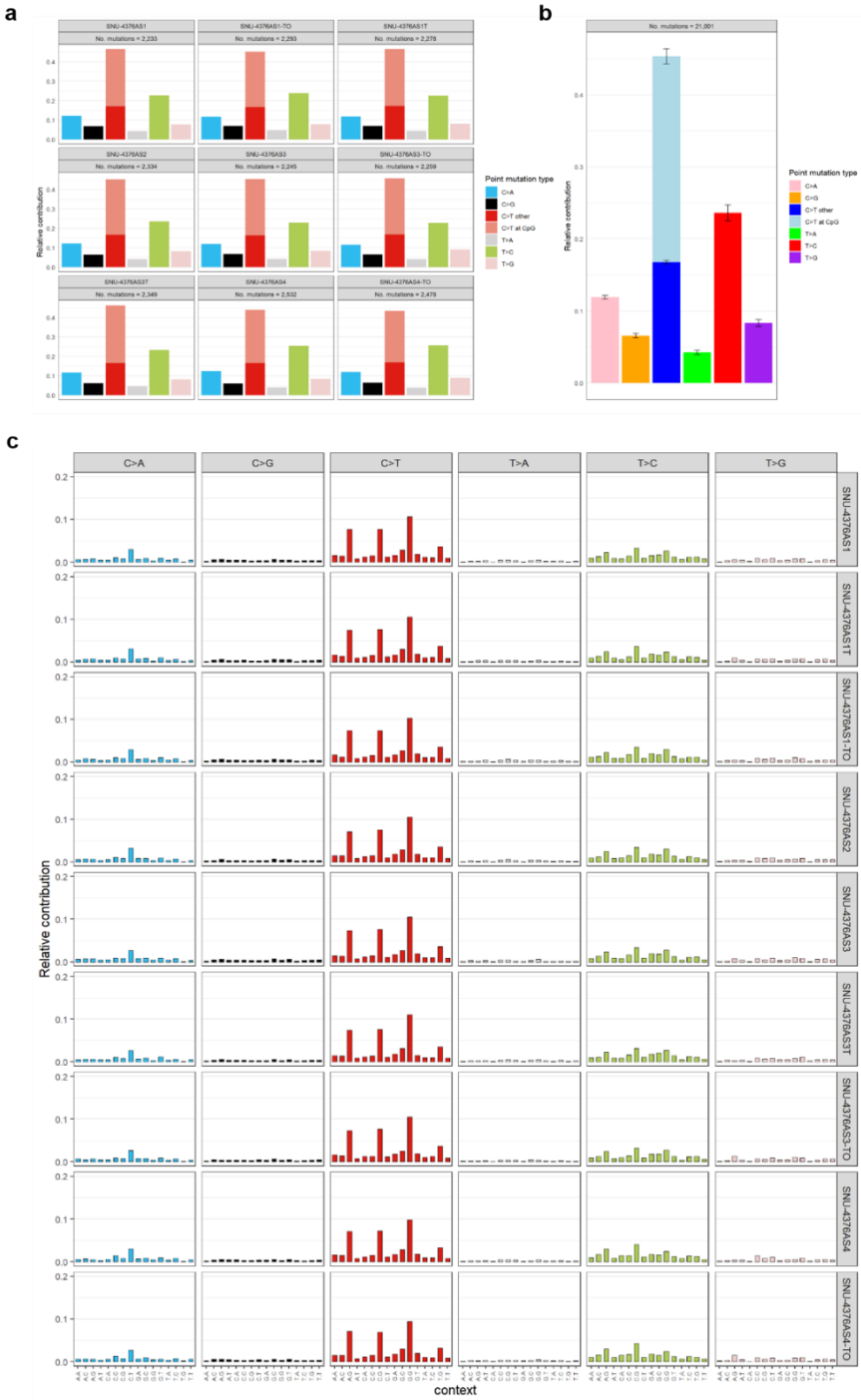
C. SNU-4351-TO


a
SNU-4351CT
No. mutations = 445
SNU-4351S1-TO
No. mutations = 494
SNU-4351S2-TO
No. mutations = 383
SNU-4351S3-TO
No. mutations = 414
SNU-4351S4-TO
No. mutations = 429
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
b
No. mutations = 2,165
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
c
C>A
C>G
C>T
T>A
T>C
T>G
SNU-4351CT
SNU-4351S1-TO
SNU-4351S2-TO
SNU-4351S3-TO
SNU-4351S4-TO
Relative contribution
context

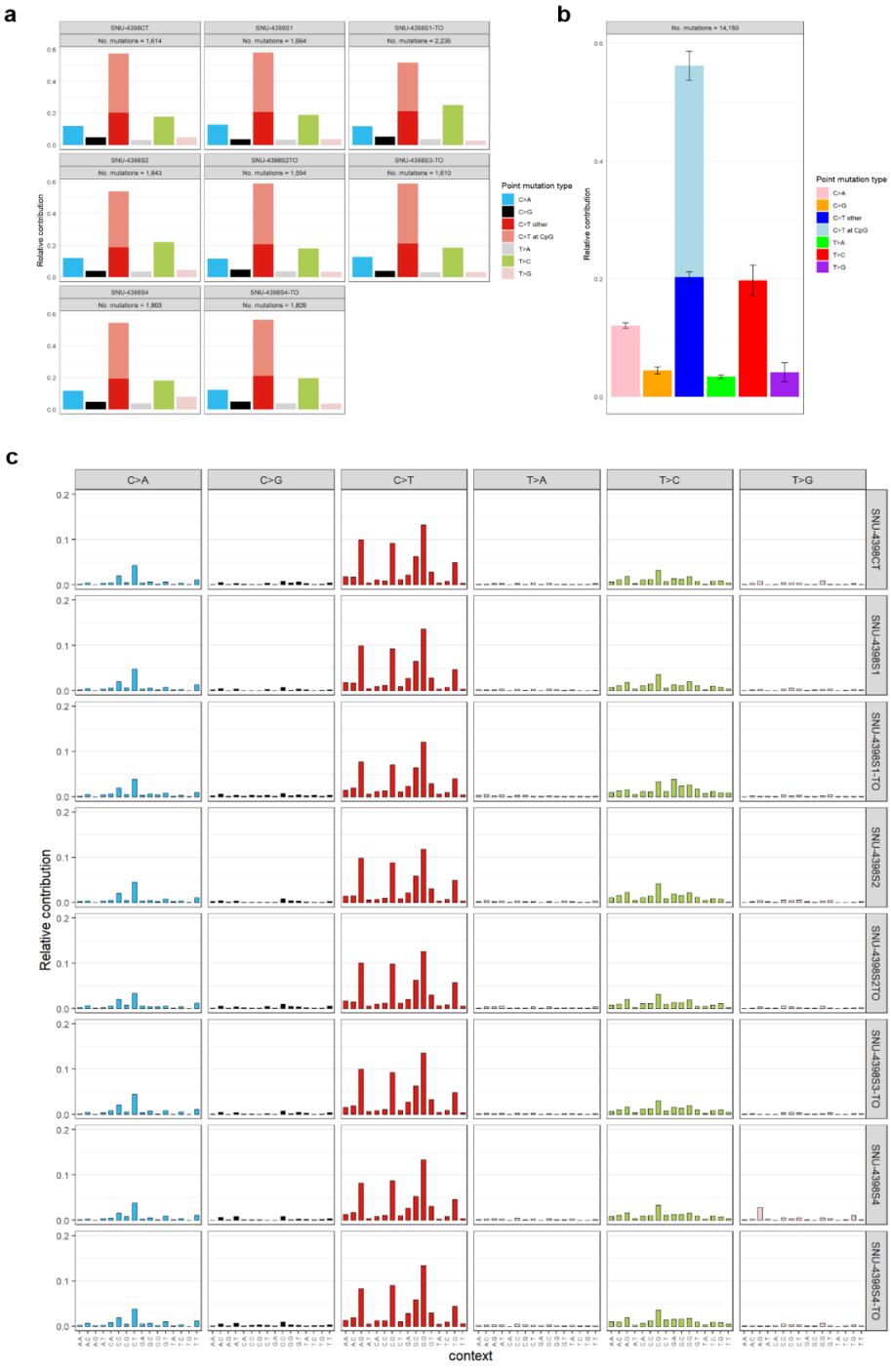
D. SNU-4374-TO


a
SNU-4374CT
No. mutations = 365
SNU-4374S1-TO
No. mutations = 395
SNU-4374S2-TO
No. mutations = 453
SNU-4374S3-TO
No. mutations = 435
SNU-4374S4-TO
No. mutations = 441
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
b
No. mutations = 2,089
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
c
C>A
C>G
C>T
T>A
T>C
T>G
SNU-4374CT
SNU-4374S1-TO
SNU-4374S2-TO
SNU-4374S3-TO
SNU-4374S4-TO
Relative contribution
context

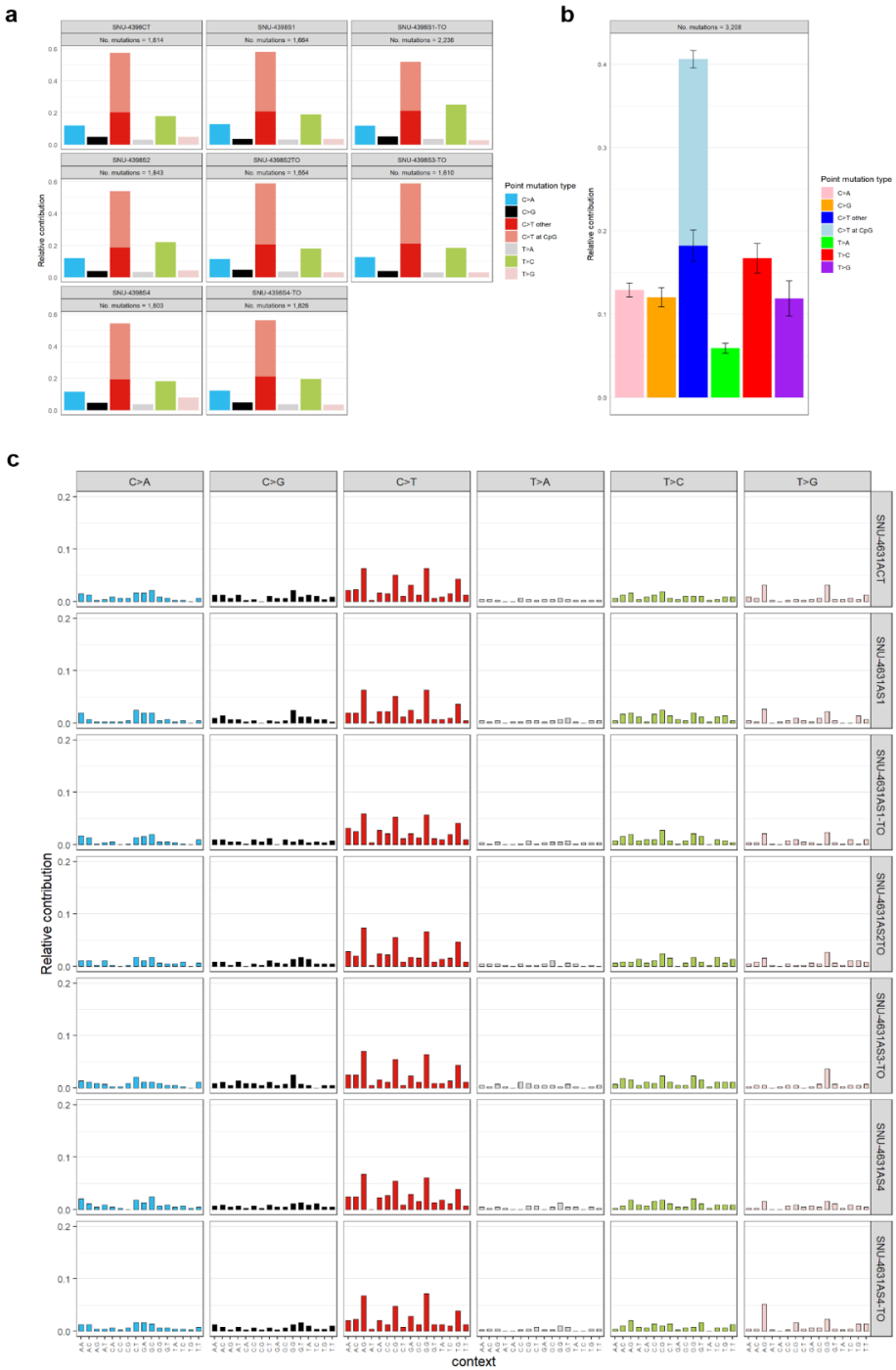
E. SNU-4376A-TO



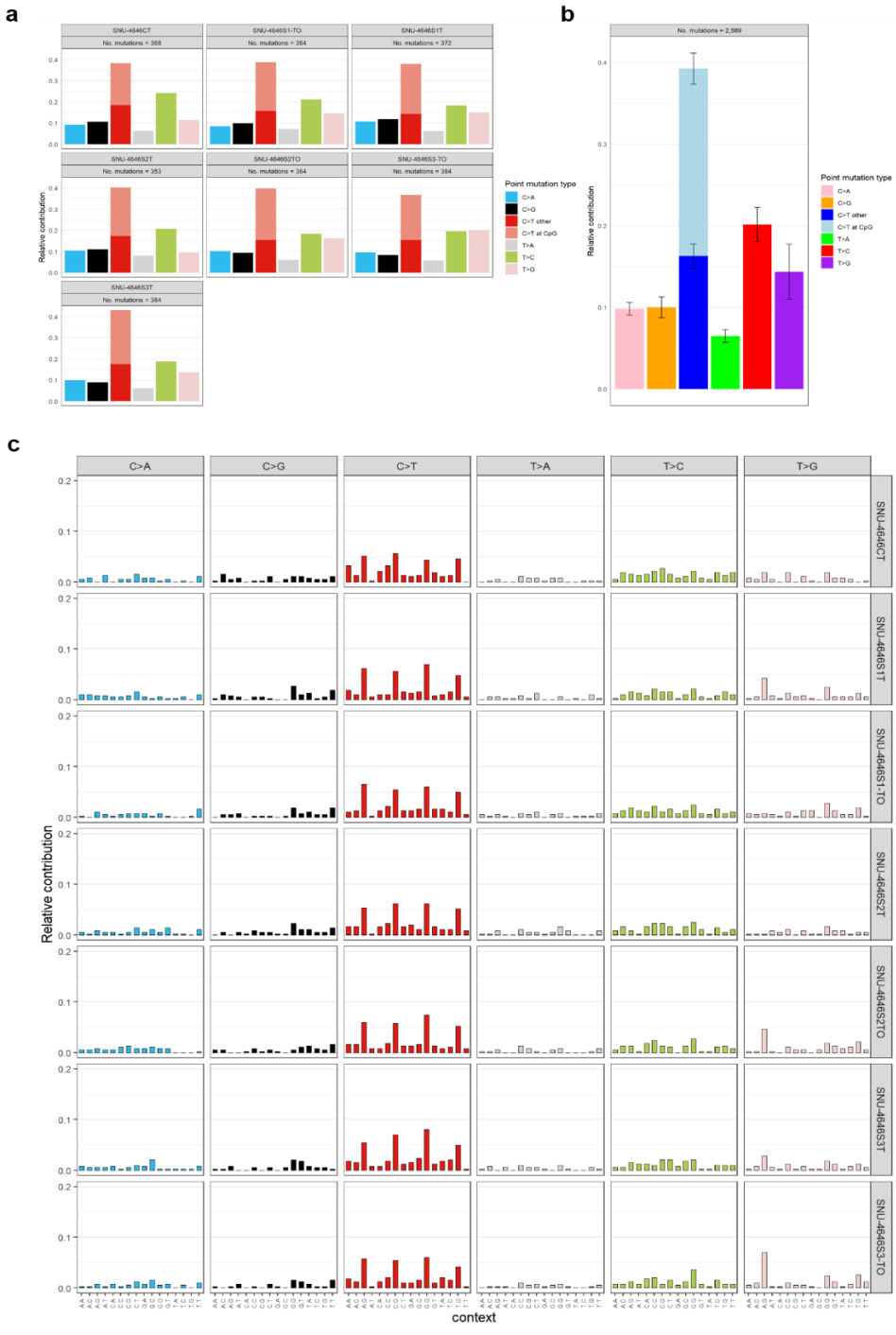
F. SNU-4398-TO


a
SNU-4398CT
No. mutations = 1,614
SNU-4398S1
No. mutations = 1,664
SNU-4398S1-TO
No. mutations = 2,236
SNU-4398S2
No. mutations = 1,843
SNU-4398S2TO
No. mutations = 1,554
SNU-4398S3-TO
No. mutations = 1,610
SNU-4398S4
No. mutations = 1,803
SNU-4398S4-TO
No. mutations = 1,826
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
b
No. mutations = 14,150
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
c
C>A
C>G
C>T
T>A
T>C
T>G
SNU-4398CT
SNU-4398S1
SNU-4398S1-TO
SNU-4398S2
SNU-4398S2TO
SNU-4398S3-TO
SNU-4398S4
SNU-4398S4-TO
Relative contribution
context

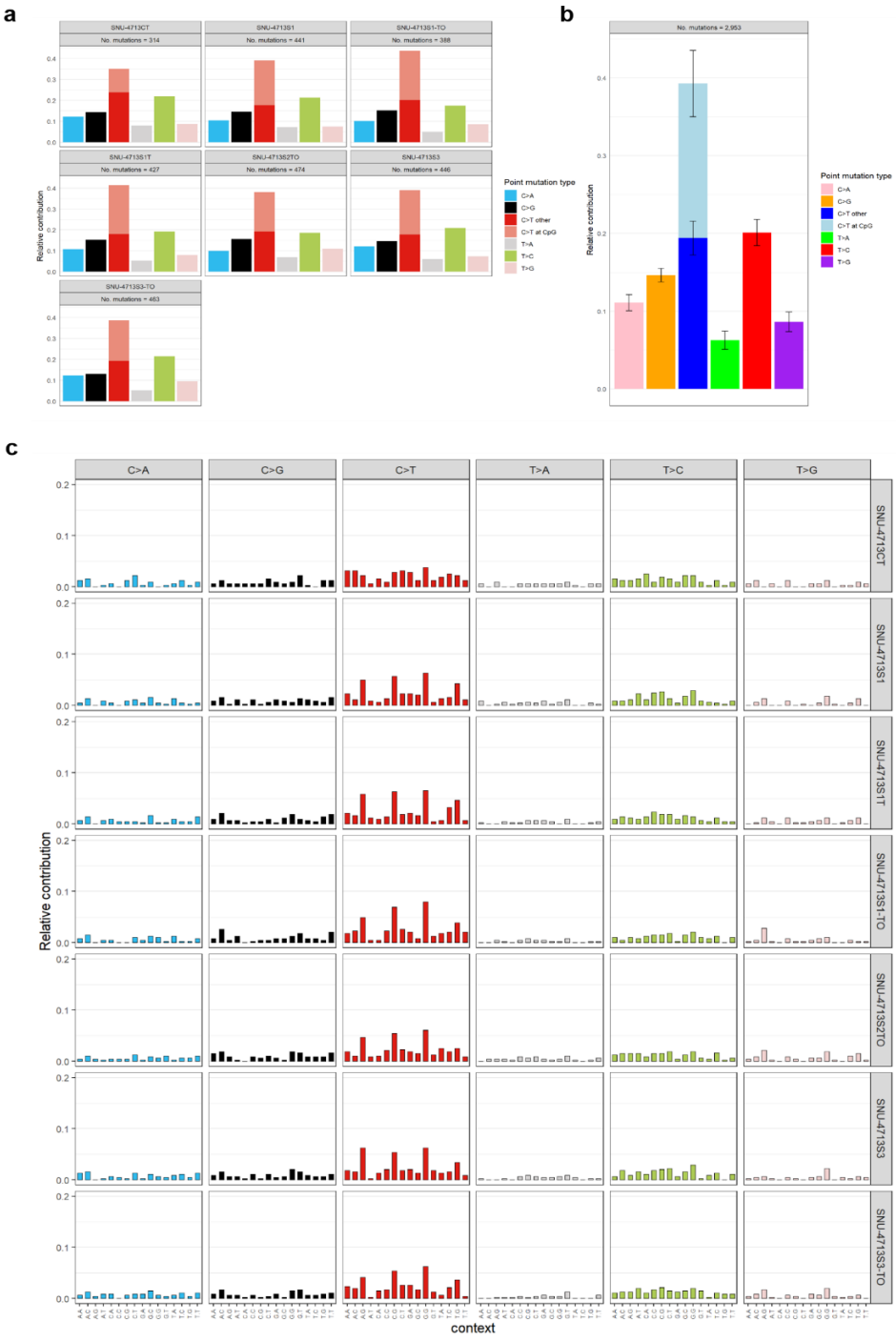
G. SNU-4631A-TO



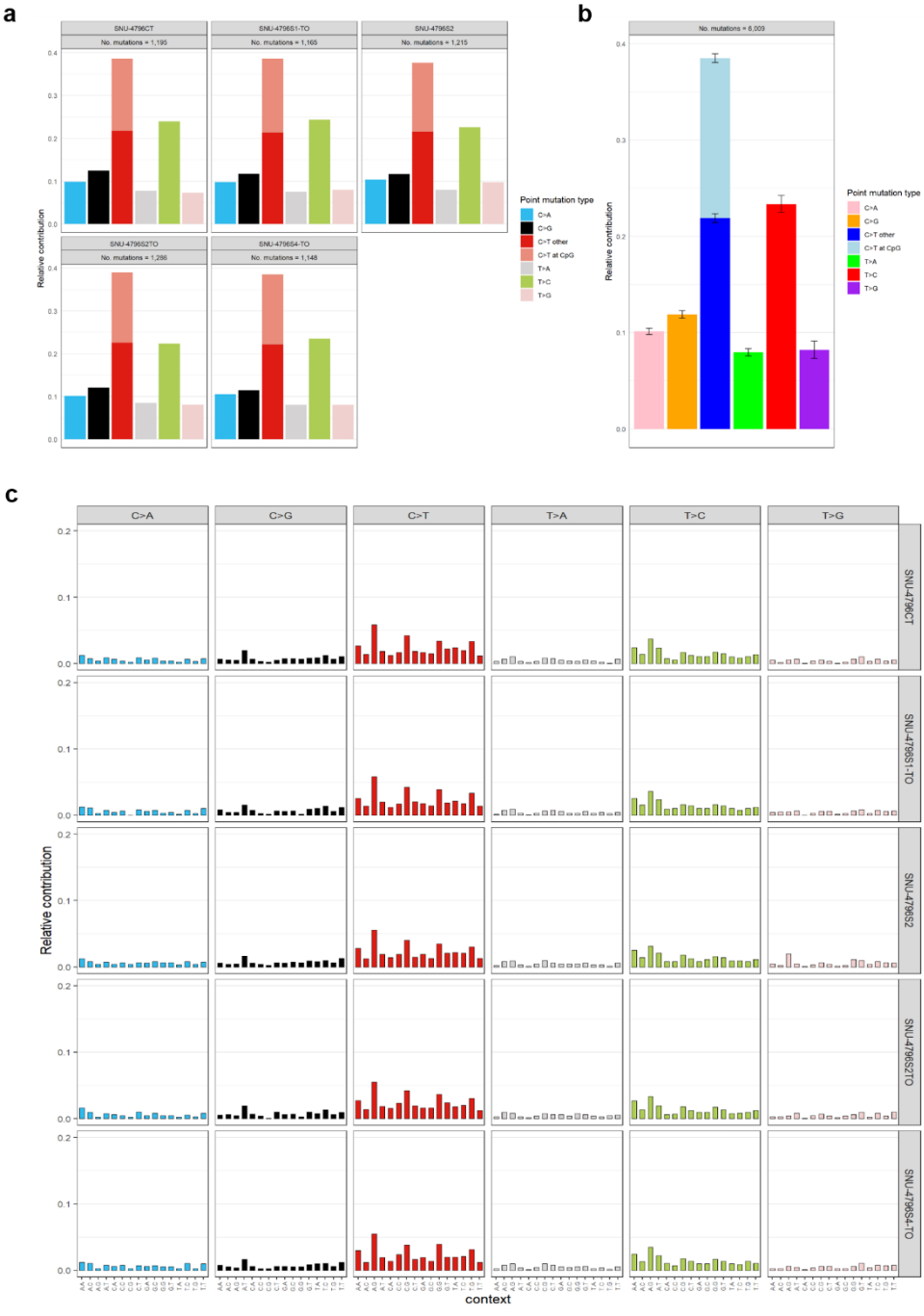
H. SNU-4646-TO


a
SNU-4646CT
No. mutations = 368
SNU-4646S1-TO
No. mutations = 364
SNU-4646S1T
No. mutations = 372
SNU-4646S2T
No. mutations = 353
SNU-4646S2TO
No. mutations = 364
SNU-4646S3-TO
No. mutations = 384
SNU-4646S3T
No. mutations = 384
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
b
No. mutations = 2,589
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
c
C>A
C>G
C>T
T>A
T>C
T>G
SNU-4646CT
SNU-4646S1T
SNU-4646S1-TO
SNU-4646S2T
SNU-4646S2TO
SNU-4646S3T
SNU-4646S3-TO
Relative contribution
context

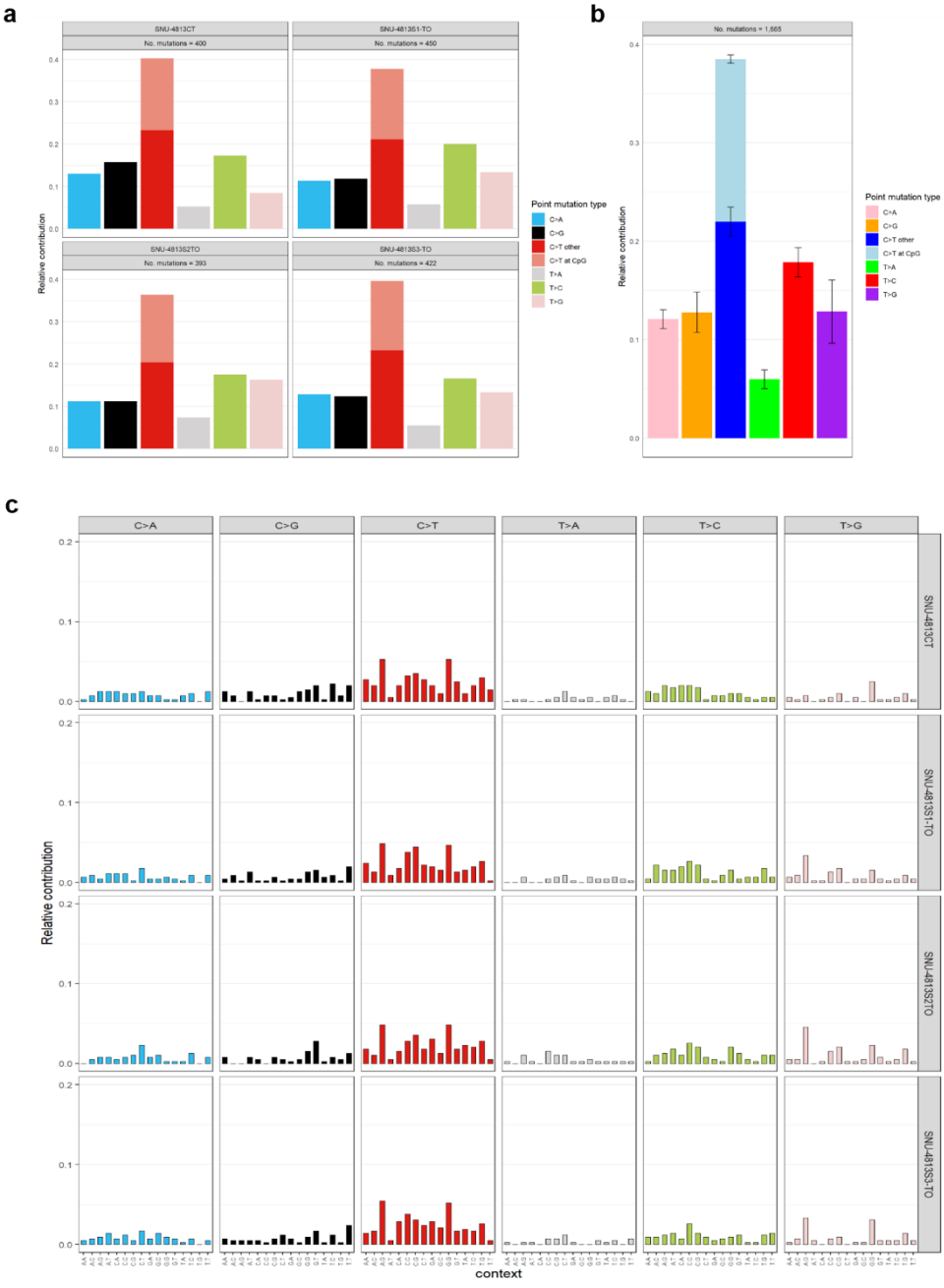
I. SNU-4713-TO


a
SNU-4713CT
No. mutations = 314
SNU-4713S1
No. mutations = 441
SNU-4713S1-TO
No. mutations = 388
SNU-4713S1T
No. mutations = 427
SNU-4713S2TO
No. mutations = 474
SNU-4713S3
No. mutations = 446
SNU-4713S3-TO
No. mutations = 463
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
b
No. mutations = 2,953
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
c
C>A
C>G
C>T
T>A
T>C
T>G
SNU-4713CT
SNU-4713S1
SNU-4713S1T
SNU-4713S1-TO
SNU-4713S2TO
SNU-4713S3
SNU-4713S3-TO
Relative contribution
context

J. SNU-4796-TO


a
SNU-4796CT
No. mutations = 1,195
SNU-4796S1-TO
No. mutations = 1,165
SNU-4796S2
No. mutations = 1,215
SNU-4796S2TO
No. mutations = 1,286
SNU-4796S4-TO
No. mutations = 1,148
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
b
No. mutations = 6,009
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
c
C>A
C>G
C>T
T>A
T>C
T>G
SNU-4796CT
SNU-4796S1-TO
SNU-4796S2
SNU-4796S2TO
SNU-4796S4-TO
Relative contribution
context

K. SNU-4813-TO


a
SNU-4813CT
No. mutations = 400
SNU-481351-TO
No. mutations = 450
SNU-481352TO
No. mutations = 393
SNU-481353-TO
No. mutations = 422
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
b
No. mutations = 1,665
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
c
C>A
C>G
C>T
T>A
T>C
T>G
SNU-4813CT
SNU-481351-TO
SNU-481352TO
SNU-481353-TO
Relative contribution
context

L. SNU-4849-TO

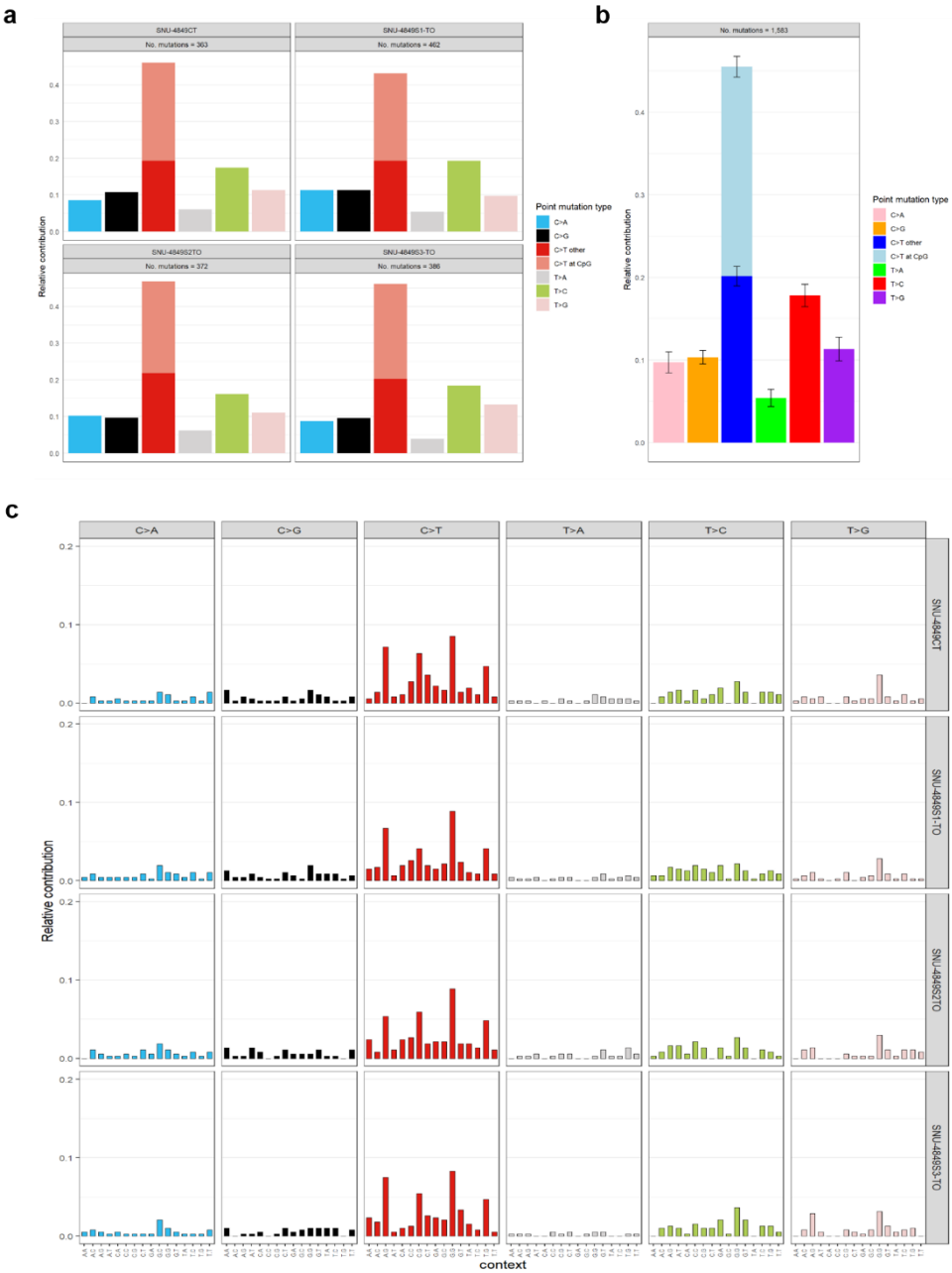

a
SNU-4849CT
No. mutations = 363
SNU-484951-TO
No. mutations = 462
SNU-484952TO
No. mutations = 372
SNU-484953-TO
No. mutations = 386
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
b
No. mutations = 1,583
Relative contribution
Point mutation type
C>A
C>G
C>T other
C>T at CpG
T>A
T>C
T>G
c
C>A
C>G
C>T
T>A
T>C
T>G
SNU-4849CT
SNU-484951-TO
SNU-484952TO
SNU-484953-TO
Relative contribution
context

Figure 6A-L. Mutational signature of twelve CRC series. A-L represent SNU-4139, SNU-4146, SNU-4351, SNU-4374, SNU-4376A, SNU-4398, SNU-4631A, SNU-4646, SNU-4713, SNU-4796, SNU-4813, and SNU-4948 series respectively.

Mutational signature of each derivate was analyzed in accordance with (a) relative contribution of point mutation type, (b) sum of relative contribution and (c) matrix of relative contribution. Cell lines and organoids that were derived from same origin displayed similar pattern in relative contribution of point mutation type.

Table 5. Mutational type occurrence of twelve CRC series

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4139CT	43	51	151	16	63	38	76	75
SNU-4139S1-TO	43	60	161	27	76	44	69	92
SNU-4139S1	51	55	163	25	67	35	85	78
SNU-4139S2-TO	43	42	138	21	55	61	65	73
SNU-4139S3-TO	47	51	164	28	65	56	70	94
SNU-4139S3	48	50	142	16	70	43	66	76
SNU-4139S4-TO	51	51	139	25	59	34	61	78

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4146CT	46	50	136	28	90	56	78	58
SNU-4146S1-TO	37	43	135	29	79	31	77	58
SNU-4146S1T	42	44	145	32	82	64	81	64
SNU-4146S2-TO	48	51	143	27	65	54	83	60
SNU-4146S2	51	46	148	22	85	50	79	69
SNU-4146S3-TO	50	54	151	27	95	42	84	67
SNU-4146S3	57	56	152	22	96	56	79	73
SNU-4146S4-TO	50	51	163	35	93	46	79	84
SNU-4146S4	47	49	145	23	81	41	76	69
SNU-4146S4T	56	49	159	26	90	62	79	80

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4351CT	46	47	193	25	99	35	104	89
SNU-4351S1-TO	66	45	213	28	93	49	104	109
SNU-4351S2-TO	57	38	164	19	71	34	91	73
SNU-4351S3-TO	57	40	184	18	78	37	98	86
SNU-4351S4-TO	70	35	187	20	78	39	101	86

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4374CT	45	55	119	25	74	47	59	60
SNU-4374S1-TO	45	53	139	28	82	48	71	68
SNU-4374S2-TO	50	57	144	35	98	69	71	73
SNU-4374S3-TO	49	45	158	38	89	56	75	83
SNU-4374S4-TO	52	57	158	34	83	57	90	68

Continued

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4376AS1-TO	270	152	1040	94	506	171	660	380
SNU-4376AS1	269	156	1059	98	514	182	668	391
SNU-4376AS1T	268	157	1034	109	545	180	652	382
SNU-4376AS2	284	151	1058	97	551	193	667	391
SNU-4376AS3-TO	268	154	1020	97	516	190	651	369
SNU-4376AS3	274	145	1082	108	548	192	695	387
SNU-4376AS3T	261	149	1036	92	515	206	656	380
SNU-4376AS4-TO	313	153	1112	99	641	214	693	419
SNU-4376AS4	295	158	1077	94	633	221	658	419

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4398CT	194	79	928	48	288	77	603	325
SNU-4398S1-TO	211	59	968	52	314	60	624	344
SNU-4398S1	264	114	1159	78	561	60	687	472
SNU-4398S2-TO	223	72	996	64	408	80	651	345
SNU-4398S2	180	74	914	56	279	51	595	319
SNU-4398S3-TO	204	63	947	50	297	49	607	340
SNU-4398S4-TO	208	84	979	65	326	141	633	346
SNU-4398S4	223	87	1027	68	356	65	645	382

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4631ACT	64	66	187	24	69	62	104	83
SNU-4631AS1-TO	53	51	153	26	78	48	88	65
SNU-4631AS1	62	55	213	31	89	57	107	106
SNU-4631AS2-TO	53	52	193	24	76	53	109	84
SNU-4631AS3-TO	57	57	174	30	81	41	102	72
SNU-4631AS4-TO	63	49	188	27	70	45	98	90
SNU-4631AS4	61	55	196	27	71	77	110	86

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4646CT	34	39	141	23	89	42	73	68
SNU-4646S1-TO	40	44	141	23	68	56	88	53
SNU-4646S1T	31	36	141	26	77	53	84	57
SNU-4646S2-TO	37	39	142	28	73	34	81	61
SNU-4646S2T	37	34	145	22	67	59	89	56
SNU-4646S3-TO	38	34	165	23	72	52	98	67
SNU-4646S3T	37	32	141	22	75	77	82	59

Continued

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4713CT	38	45	110	25	69	27	35	75
SNU-4713S1-TO	46	64	172	32	94	33	94	78
SNU-4713S1	46	65	177	23	82	34	100	77
SNU-4713S1T	39	59	170	19	68	33	92	78
SNU-4713S2-TO	47	74	180	33	88	52	89	91
SNU-4713S3-TO	54	65	174	27	93	33	95	79
SNU-4713S3	57	60	179	24	99	44	90	89

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4796CT	118	149	462	93	286	87	202	260
SNU-4796S1-TO	114	137	450	87	284	93	201	249
SNU-4796S2-TO	126	142	458	97	274	118	196	262
SNU-4796S2	130	155	501	109	287	104	211	290
SNU-4796S4-TO	121	131	442	92	270	92	188	254

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4813CT	52	63	161	21	69	34	68	93
SNU-4813S1-TO	51	53	170	26	90	60	75	95
SNU-4813S2-TO	44	44	143	29	69	64	63	80
SNU-4813S3-TO	54	52	167	23	70	56	69	98

SampleCode	C>A	C>G	C>T	T>A	T>C	T>G	C>T at CpG	C>T other
SNU-4849CT	31	39	167	22	63	41	97	70
SNU-4849S1-TO	52	52	199	25	89	45	110	89
SNU-4849S2-TO	38	36	174	23	60	41	93	81
SNU-4849S3-TO	34	37	178	15	71	51	100	78

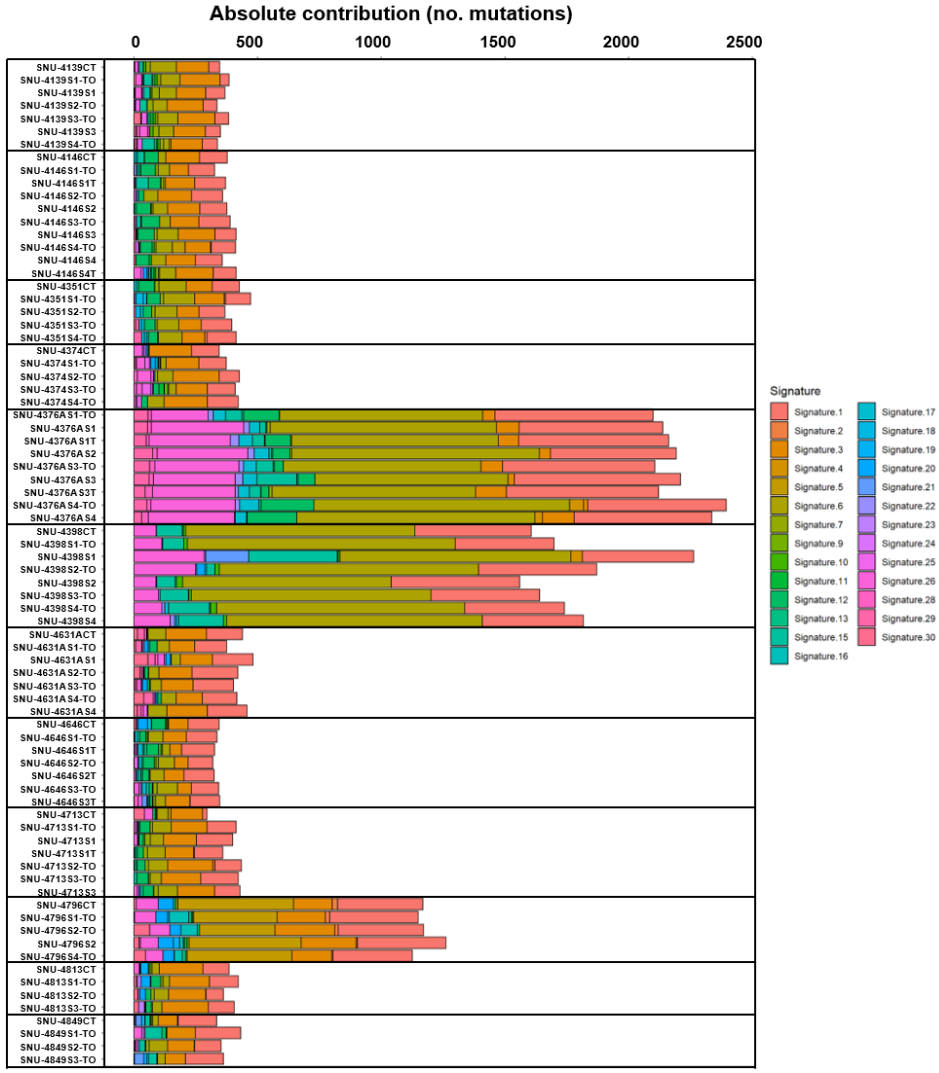


Figure 7. Total mutational load and mutational signatures of colorectal cancer cell lines/organoids and paired primary tumors. Different colors represent 30 kinds of signatures

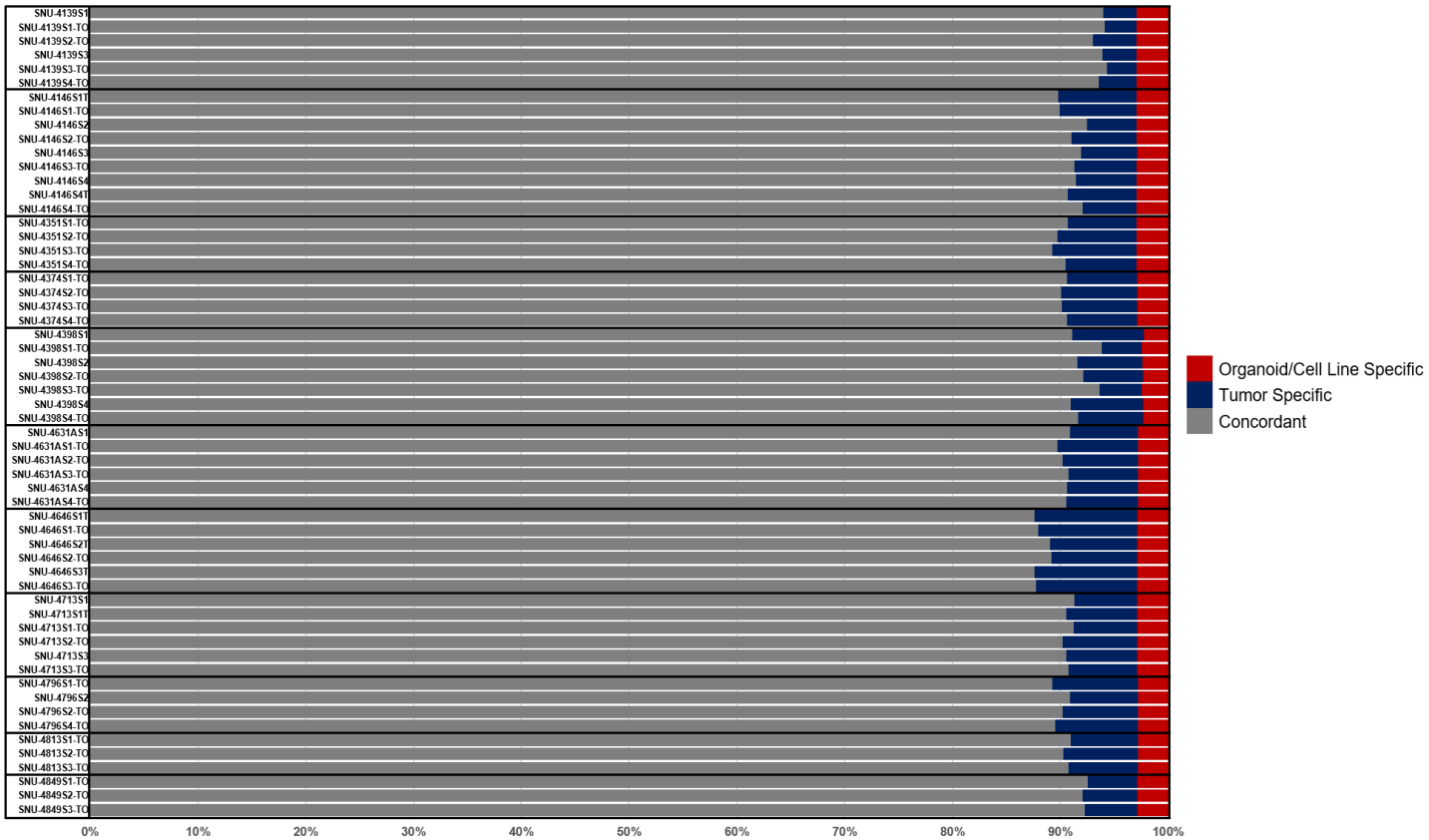

SNU-4139S1
SNU-4139S1-TC
SNU-4139S2-TC
SNU-4139S2
SNU-4139S3-TC
SNU-4139S4-TC
SNU-4146S1T
SNU-4146S1-TC
SNU-4146S2
SNU-4146S2-TC
SNU-4146S3
SNU-4146S3-TC
SNU-4146S4
SNU-4146S4T
SNU-4146S4-TC
SNU-4351S1-TC
SNU-4351S2-TC
SNU-4351S3-TC
SNU-4351S4-TC
SNU-4374S1-TC
SNU-4374S2-TC
SNU-4374S3-TC
SNU-4374S4-TC
SNU-4398S1
SNU-4398S1-TC
SNU-4398S2
SNU-4398S2-TC
SNU-4398S3-TC
SNU-4398S4
SNU-4398S4-TC
SNU-4631AS1
SNU-4631AS1-TC
SNU-4631AS2-TC
SNU-4631AS3-TC
SNU-4631AS4
SNU-4631AS4-TC
SNU-4646S1T
SNU-4646S1-TC
SNU-4646S2T
SNU-4646S2-TC
SNU-4646S3T
SNU-4646S3-TC
SNU-4713S1
SNU-4713S1T
SNU-4713S1-TC
SNU-4713S2-TC
SNU-4713S3
SNU-4713S3-TC
SNU-4796S1-TC
SNU-4796S2
SNU-4796S2-TC
SNU-4796S4-TC
SNU-4813S1-TC
SNU-4813S2-TC
SNU-4813S3-TC
SNU-4849S1-TC
SNU-4849S2-TC
SNU-4849S3-TC
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Organoid/Cell Line Specific
Tumor Specific
Concordant

Figure 8. Histogram showing the concordance (percent) of SNVs between colorectal cancer cell lines/organoids and corresponding primary tumors.

Mutational concordance within the coding regions in both tumor cell lines and organoids was highly corresponded with the matched tumor specimen for both hypermutated and non-hypermutated patients (median = 0.90 frequency of concordance, range 0.87 to 0.94). Although there existed a slight variation within a same series, the difference between samples were imperceptible.

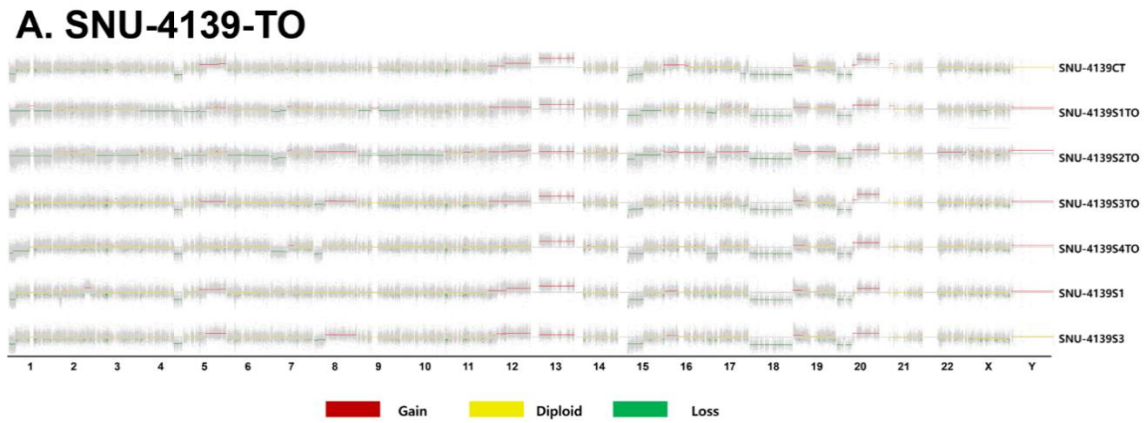
Table 6. Mutational Concordance between colorectal cancer cell lines/organoids and corresponding primary tumors

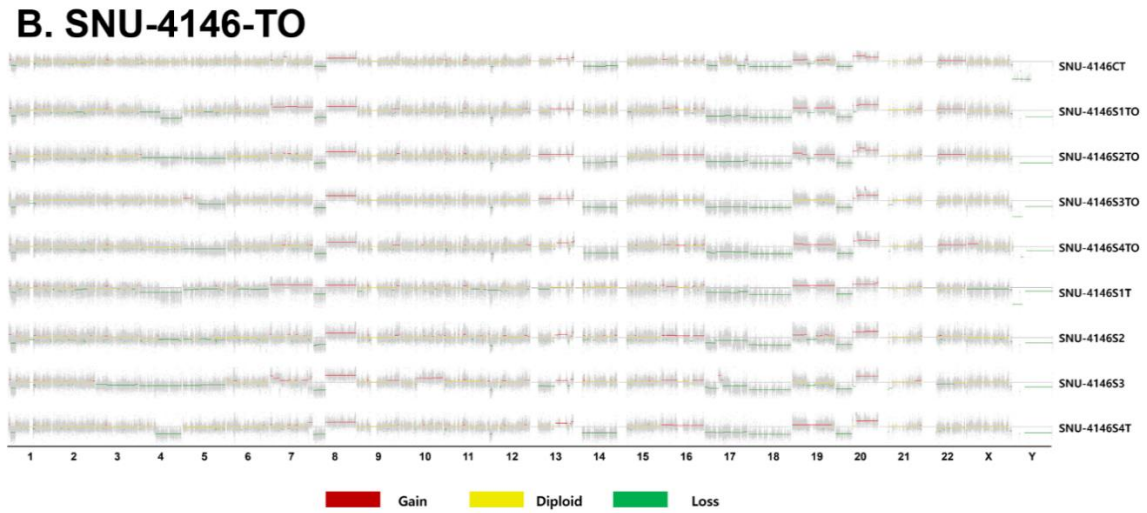
SAMPLECODE	SampleOnly	TumorOnly	Shared	Total	Sample Only (%)	Tumor Only (%)	Concordant (%)
SNU-4139S1	372	387	11898	12657	2.939	3.058	94.003
SNU-4139S1-TO	415	375	11910	12700	2.929	2.953	94.118
SNU-4139S2-TO	369	509	11776	12654	2.940	4.022	93.038
SNU-4139S3	348	395	11890	12633	2.945	3.127	93.929
SNU-4139S3-TO	370	342	11943	12655	2.940	2.702	94.358
SNU-4139S4-TO	333	436	11849	12618	2.948	3.455	93.596
SNU-4146S1T	2015	927	9775	12717	2.925	7.289	89.785
SNU-4146S1-TO	1965	905	9797	12667	2.937	7.145	89.919
SNU-4146S2	2018	588	10114	12720	2.925	4.623	92.453
SNU-4146S2-TO	1991	762	9940	12693	2.931	6.003	91.066
SNU-4146S3	2184	665	10037	12886	2.887	5.161	91.953
SNU-4146S3-TO	2024	737	9966	12727	2.923	5.791	91.286
SNU-4146S4	1983	710	9993	12686	2.932	5.597	91.471
SNU-4146S4T	2049	812	9891	12752	2.917	6.368	90.715
SNU-4146S4-TO	2014	636	10067	12717	2.925	5.001	92.074
SNU-4351S1-TO	670	817	11336	12823	2.901	6.371	90.728
SNU-4351S2-TO	546	931	11222	12699	2.929	7.331	89.739
SNU-4351S3-TO	560	994	11159	12713	2.926	7.819	89.255
SNU-4351S4-TO	622	841	11312	12775	2.912	6.583	90.505
SNU-4374S1-TO	382	840	11708	12930	2.877	6.497	90.626
SNU-4374S2-TO	447	915	11633	12995	2.863	7.041	90.096
SNU-4374S3-TO	386	904	11644	12934	2.876	6.989	90.135
SNU-4374S4-TO	429	845	11703	12977	2.867	6.512	90.622
SNU-4398S1	1633	1083	13700	16416	2.266	6.597	91.137
SNU-4398S1-TO	586	573	14210	15369	2.420	3.728	93.851
SNU-4398S2	1003	958	13826	15787	2.356	6.068	91.575
SNU-4398S2-TO	1326	893	13890	16109	2.309	5.543	92.147
SNU-4398S3-TO	602	607	14176	15385	2.418	3.945	93.637
SNU-4398S4	1143	1066	13717	15926	2.336	6.693	90.971
SNU-4398S4-TO	1089	950	13833	15872	2.344	5.985	91.671

Continued

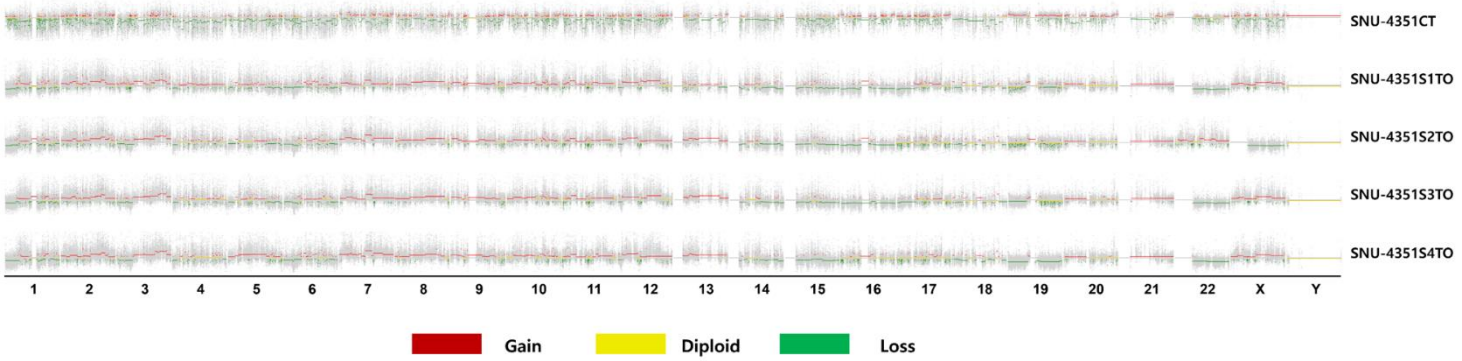
SAMPLECODE	SampleOnly	TumorOnly	Shared	Total	Sample Only (%)	Tumor Only (%)	Concordant (%)
SNU-4631AS1	412	856	12178	13446	2.767	6.366	90.867
SNU-4631AS1-TO	337	1000	12034	13371	2.782	7.479	89.739
SNU-4631AS2-TO	422	943	12091	13456	2.765	7.008	90.227
SNU-4631AS3-TO	376	870	12164	13410	2.774	6.488	90.738
SNU-4631AS4	385	882	12152	13419	2.772	6.573	90.655
SNU-4631AS4-TO	370	892	12142	13404	2.775	6.655	90.570
SNU-4646S1T	317	1246	11510	13073	2.846	9.531	87.623
SNU-4646S1-TO	342	1208	11548	13098	2.840	9.223	87.937
SNU-4646S2T	308	1061	11695	13064	2.848	8.122	89.031
SNU-4646S2-TO	333	1039	11717	13089	2.842	7.938	89.220
SNU-4646S3T	321	1252	11504	13077	2.845	9.574	87.581
SNU-4646S3-TO	360	1238	11518	13116	2.836	9.439	87.725
SNU-4713S1	514	765	11843	13122	2.835	5.830	91.335
SNU-4713S1T	478	860	11748	13086	2.843	6.572	90.585
SNU-4713S1-TO	495	773	11836	13104	2.839	5.899	91.262
SNU-4713S2-TO	530	917	11691	13138	2.831	6.980	90.189
SNU-4713S3	526	871	11737	13134	2.832	6.632	90.536
SNU-4713S3-TO	515	843	11765	13123	2.835	6.424	90.741
SNU-4796S1-TO	398	1046	11748	13192	2.820	7.929	89.251
SNU-4796S2	506	842	11952	13300	2.797	6.331	90.872
SNU-4796S2-TO	440	924	11870	13234	2.811	6.982	90.207
SNU-4796S4-TO	385	1004	11791	13180	2.822	7.618	89.560
SNU-4813S1-TO	368	828	12048	13244	2.809	6.252	90.939
SNU-4813S2-TO	344	911	11965	13220	2.814	6.891	90.295
SNU-4813S3-TO	346	851	12025	13222	2.813	6.436	90.750
SNU-4849S1-TO	414	586	11893	12893	2.885	4.545	92.570
SNU-4849S2-TO	373	644	11835	12852	2.894	5.011	92.095
SNU-4849S3-TO	379	625	11854	12858	2.893	4.861	92.246

Exome-wide CNVs of cell lines/organoids and matched tumor tissues were compared to confirm that the patterns of CNVs were maintained throughout the whole exome except for SNU-4351 series and SNU-4796 series in which primary tumor specimens and matched normal samples were of insufficient purity to determine CNVs (Figure 9). Our samples displayed comparable somatic copy number alterations (SCNAs) with much larger clinical cohort (Figure 10A) (19). Inspection of the top regions identified by TCGA disclosed the presence of ERBB2-, MYC- and BRCA2-amplified and SMAD4-depleted cell lines/organoids, as well as a documented gain of 13q in the non-hypermutated group (Figure 10B). Overall, these data validated that cell line/organoid cultures recapitulate the genomic characters of the primary tumor and most of the genomic diversity of CRC.

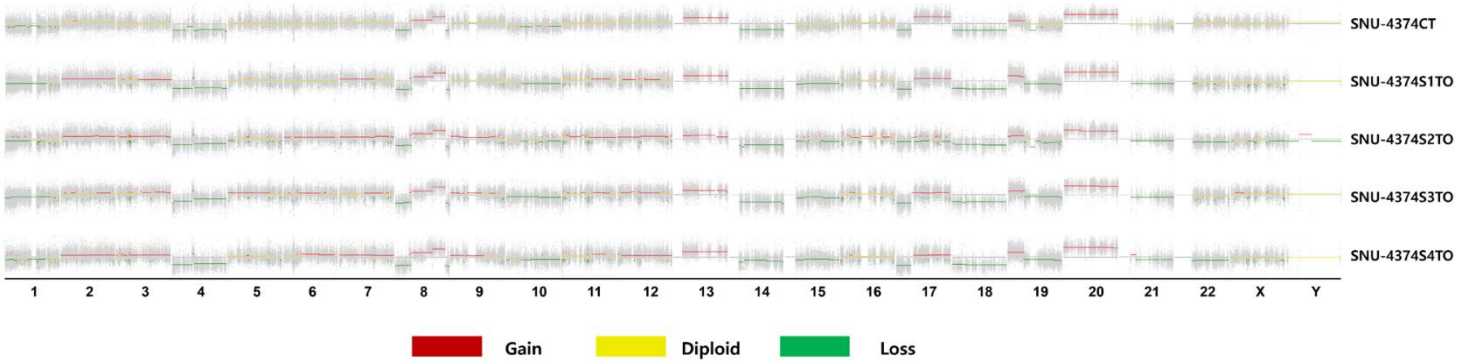

A. SNU-4139-TO
SNU-4139CT
SNU-4139S1TO
SNU-4139S2TO
SNU-4139S3TO
SNU-4139S4TO
SNU-4139S1
SNU-4139S3
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
Gain
Diploid
Loss


B. SNU-4146-TO
SNU-4146CT
SNU-4146S1TO
SNU-4146S2TO
SNU-4146S3TO
SNU-4146S4TO
SNU-4146S1T
SNU-4146S2
SNU-4146S3
SNU-4146S4T
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
Gain
Diploid
Loss

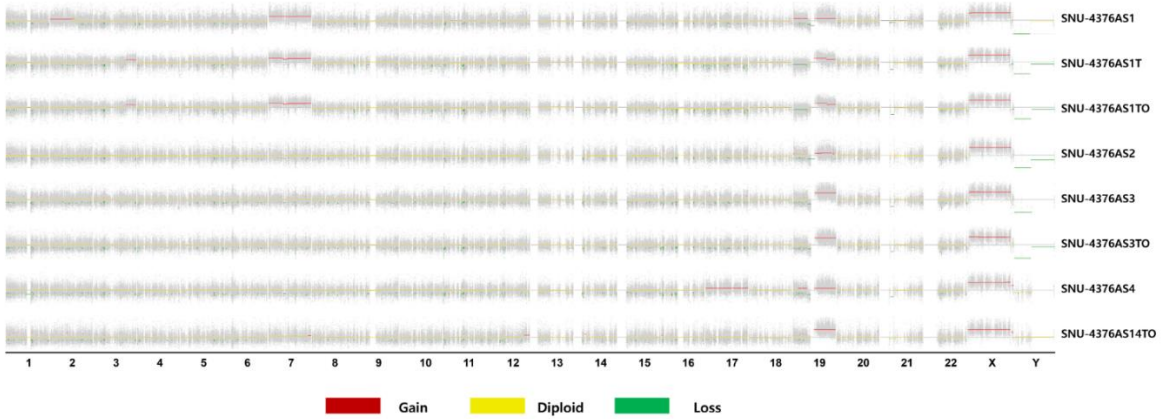
C. SNU-4351-TO



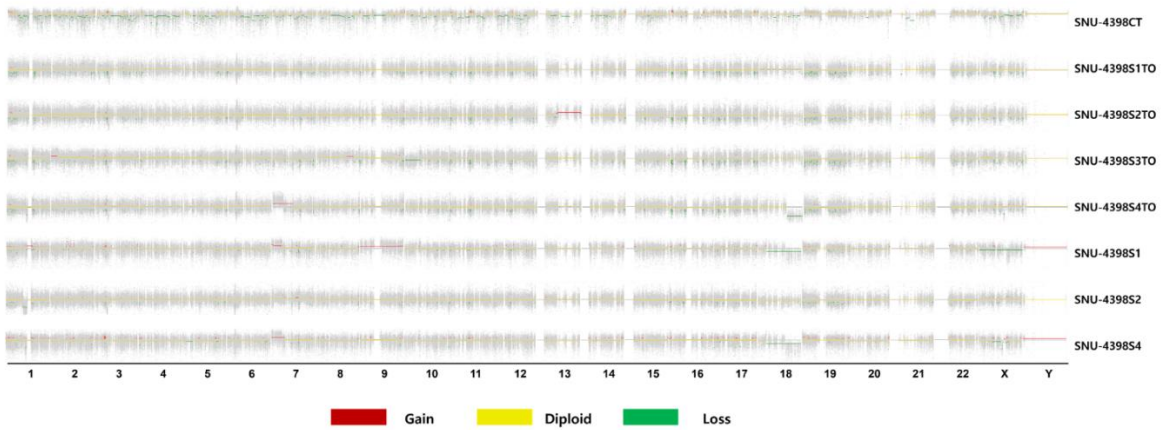
D. SNU-4374-TO



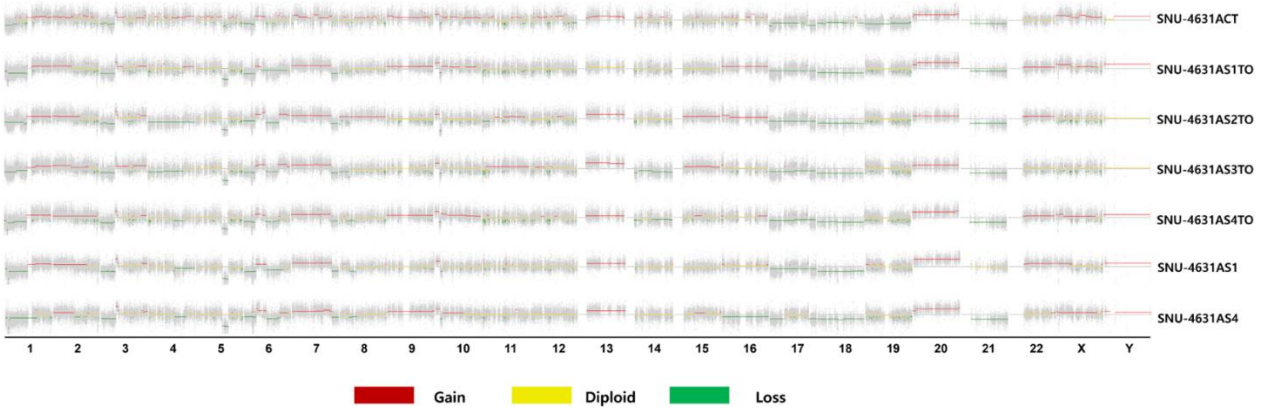
E. SNU-4376A-TO



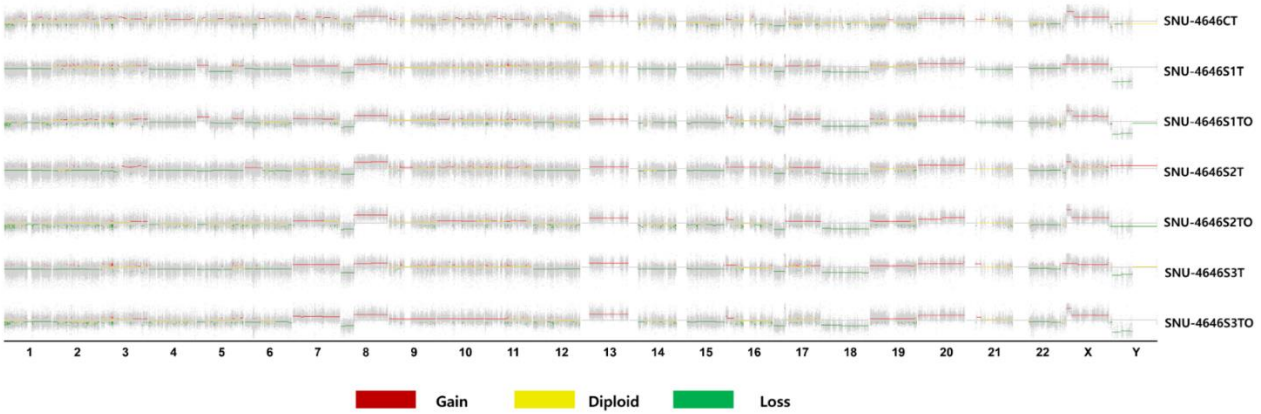
F. SNU-4398-TO



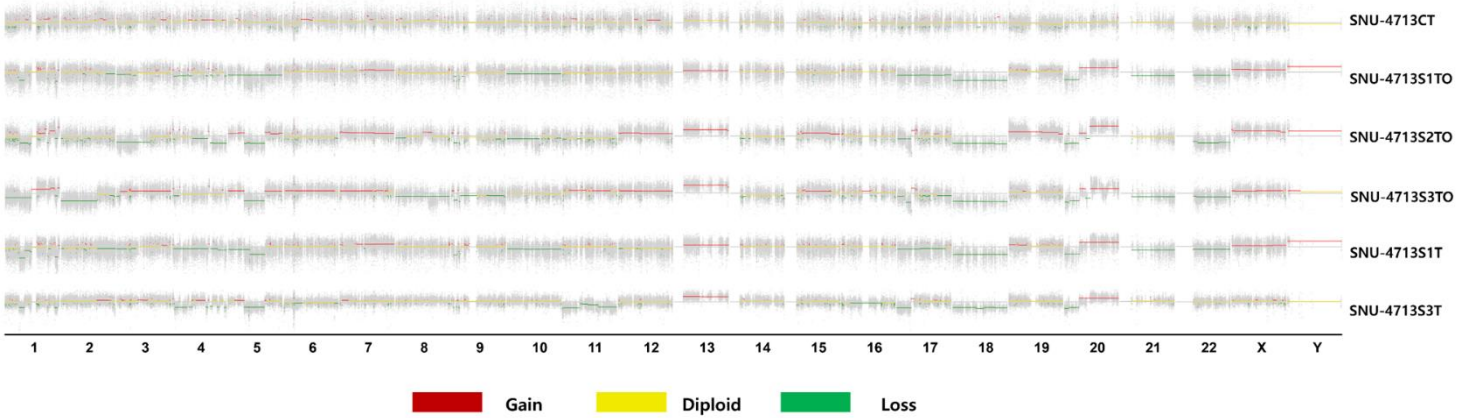
G. SNU-4631A-TO



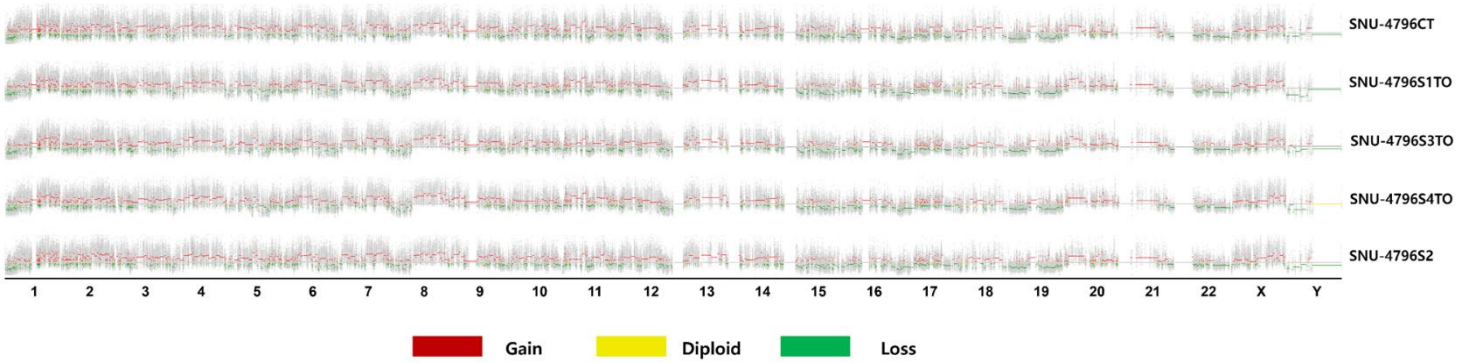
H. SNU-4646-TO



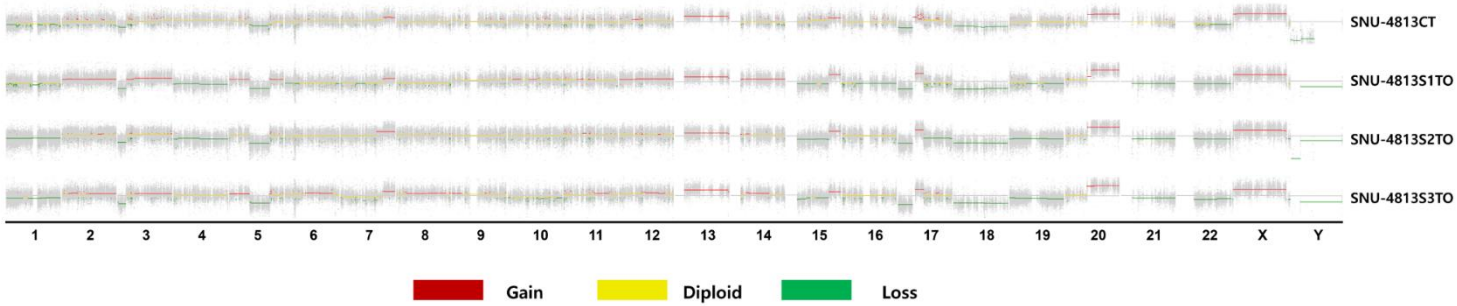
I. SNU-4713-TO



J. SNU-4796-TO



K. SNU-4813-TO



L. SNU-4849-TO

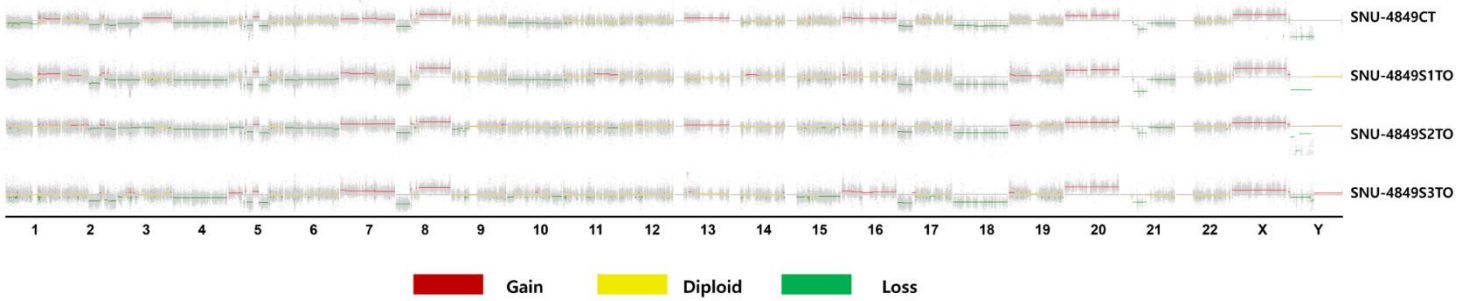


Figure 9A–L. Genome–wide gene copy number variations (CNVs) of CRC cell lines/organoids and paired primary tumors (red, gains; green, losses; yellow, diploid). A–L represent SNU–4139, SNU–4146, SNU–4351, SNU–4374, SNU–4376A, SNU–4398, SNU–4631A, SNU–4646, SNU–4713, SNU–4796, SNU–4813, and SNU–4948 series respectively. Primary tumor specimens and matched normal samples of SNU–4351_SET and SNU–4796_SET were of insufficient purity to reveal CNVs.

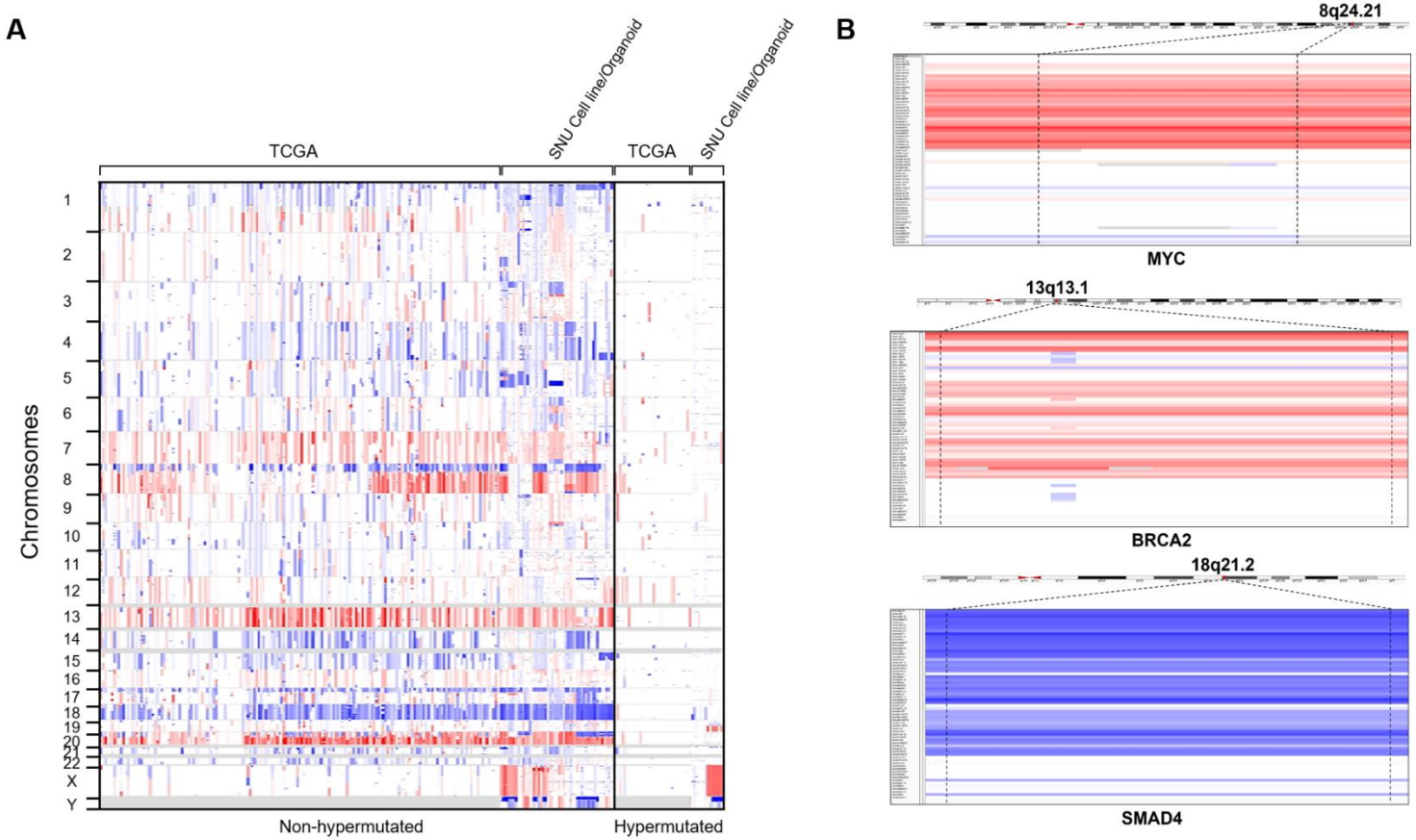

A
TCGA
SNU Cell line/Organoid
TCGA
SNU Cell line/Organoid
Chromosomes
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Non-hypermutated
Hypermutated
B
8q24.21
MYC
13q13.1
BRCA2
18q21.2
SMAD4

Figure 10A. Comparison of somatic copy number alterations found in the primary cancer tissues and corresponding cell lines/organoids (CT/Org) and TCGA CRC in both hypermutated and non-hypermutated samples. 7B. Somatic copy number alterations in organoids among frequently amplified genes identified in TCGA COAD.

Our samples displayed comparable somatic copy number alterations (SCNAs) with much larger clinical cohort. Manual inspection of the top regions identified by TCGA did reveal the presence of ERBB2-, MYC- and BRCA2-amplified and SMAD4-depleted organoids, as well as a reported gain of 13q in the non-hypermutated group.

Evolutionary histories of twelve multisampling CRC

Treeomics algorithm was applied to the multi-region sequencing data to draw evolutionary trees of the twelve CRCs (30). We comprised organoids as well as cell lines derivate to draw phylogenetic trees to determine if the culture method affects mutational aspects. Possible sequencing artifacts were rectified by Treeomics algorithm to make mutational patterns of each sample compatible with the topological variant of the evolutionary tree. Based on multi-region WES profiles, Treeomics classified somatic mutations as all (trunk), more than two samples (shared) and a single samples (individual). Due to significant differences between the number of mutations in MSS and MSI tumor derivate, the length of the trunk and braches represented the number of trunk and branch mutations respectively only within a single set. Besides, our samples involves organoid-derived 2D cell lines (its nomenclature ends with -T). This could potentially cause biased number of shared mutations between organoid-derived 2D cell lines and its parental organoids. Therefore, we also limited key mutations in drawing phylogenetic tree to known cancer driver genes reported from the PanCancer

analysis of 10,000 TCGA tumor samples (31). We consider a phylogenetic tree as long-trunk trees (SNU-4376A, -4646, -4796, -4813, and -4849 sets) when the number of trunk mutations outnumbers the sum of shared and individual mutations. Otherwise, a phylogenetic tree is considered as short-trunk trees (SNU-4139, -4146, -4351, -4374, -4398, -4631A, and -4713 sets). Interestingly, SNU-4398 series was derivate of MSI tumor but still classified as short-trunk trees. Common driver genes with potential functional mutations which includes non-synonymous single-nucleotide variants (SNVs), stopgain SNVs, splicing SNVs, or insertion/deletions (indels) were plotted for analyzing evolutionary history of each tumor (Figure 11).

For instance, SNU-4849 harbored multiple known driver mutations such as APC (c.4132C>T/ p.Q1378*), TP53 (c.713G>A / p.C238Y) and ARID1A (c.4187_4188del/ p.G1396Afs*48) with an allele frequency of ~0.95 in the trunk while each subclones shaped the phylogenetic tree with individual mutations with VAFs of ~ 0.2. This implies that the first hit mainly contributes the tumorigenesis of colon epithelial cells and thereafter the tumor had not been progressed

expressively. SNU-4374 had similar shape with SNU-4849, in which no shared mutation was observed. SNU-4374 series also had several driver mutations such as APC (c.4132C>T/ p.Q1378*), TP53 (c.406C>G/ p.Q136E) with VAFs of ~0.95 in the trunk while each subclones shaped the phylogenetic tree with individual mutations with VAFs of ~ 0.3. Nevertheless, S2TO clone had protruding acquisition of mutational burden in KMT2C and SOX9 genes with VAFs of ~0.35, which suggested one major branch was forming with epigenetic factors (Figure 11 and Figure 12).

The original tumor tissue of SNU-4146 had TP53 mutation (c.817C>T/ p.R273C) with VAFs of 1, which appeared in the relatively early stage of evolution and two APC mutations (c.637C>T/p.R213* and c.1312+2T>G/p.X438_splice) with VAFs of 0.51 and 0.43 respectively. We observed that VAFs of nonsense APC mutation (c.637C>T/p.R213*) were ascended to 0.71 and 0.82 in subclones S2TO and S3TO respectively whereas VAFs of other two subclones remained unchanged. In contrast, splice site APC mutation (c.1312+2T>G/p.X438_splice) had decreased VAFs of 0.37 and 0.26 in subclones S2TO and S3TO respectively while VAFs of other two

subclones remained unchanged as well. Even though two APC mutations were present in all subclones, this analysis reflected two major subclones were subjected to loss of heterozygosity (LOH) leading to biallelic inactivation of APC. Notably, this tendency was reversed in 2D culture cell lines, which could indicate different culture methods favor particular type of mutations (Figure 11 and Figure 12).

Two MSI tumor derivative series (SNU-4376A and SNU-4398 sets) shared ~50% of somatic mutations. They displayed low VAFs of major hit mutations including APC, TP53, PIK3CA, and KRAS, and discordant driving alterations in APC and TP53 existed, suggesting hyper-mutated phenotype may have been present prior to the LOH leading to biallelic inactivation of APC and TP53 and acquisition of growth promoting mutations. Other ten cases share at least one major hit mutations with VAFs of ~0.95, which indicated identical somatically mutated progenitor cell was mutually ancestral but then diverged to acquire independent secondary alterations. Comparisons between physical locations of sub-clones in the total tumor mass and the constructed evolutionary tree indicated that sub-clonal

dispersion normally progressed in spatially associated ways (Figure 11 and Figure 12).


SNU-4139 SET
SNU-4146 SET
SNU-4351 SET
SNU-4374 SET
SNU-4376A SET
SNU-4398 SET
SNU-4631A SET
SNU-4646 SET
SNU-4713 SET
SNU-4796 SET
SNU-4813 SET
SNU-4849 SET

Figure 11. Evolutionary trees of twelve CRC series. Using Treeomics algorithm, phylogenetic trees were drawn from the multi-regional WES data. Based on multi-region WES profiles, Treeomics classified somatic mutations as all (trunk), more than two samples (shared) and a single samples (individual). Leaves match to each sub-clonal samples. The total number of mutations indicated by numbers adjacent to the roots of each branches correspond to lengths of the trunk and branches within a samples. Possible driver mutations that are listed in Cancer Gene Census List are mapped along the phylogenetic trees. The actual picture of each tumor mass is presented with locations where each sub-clonal sample was obtained. White scale bars = 1 cm.

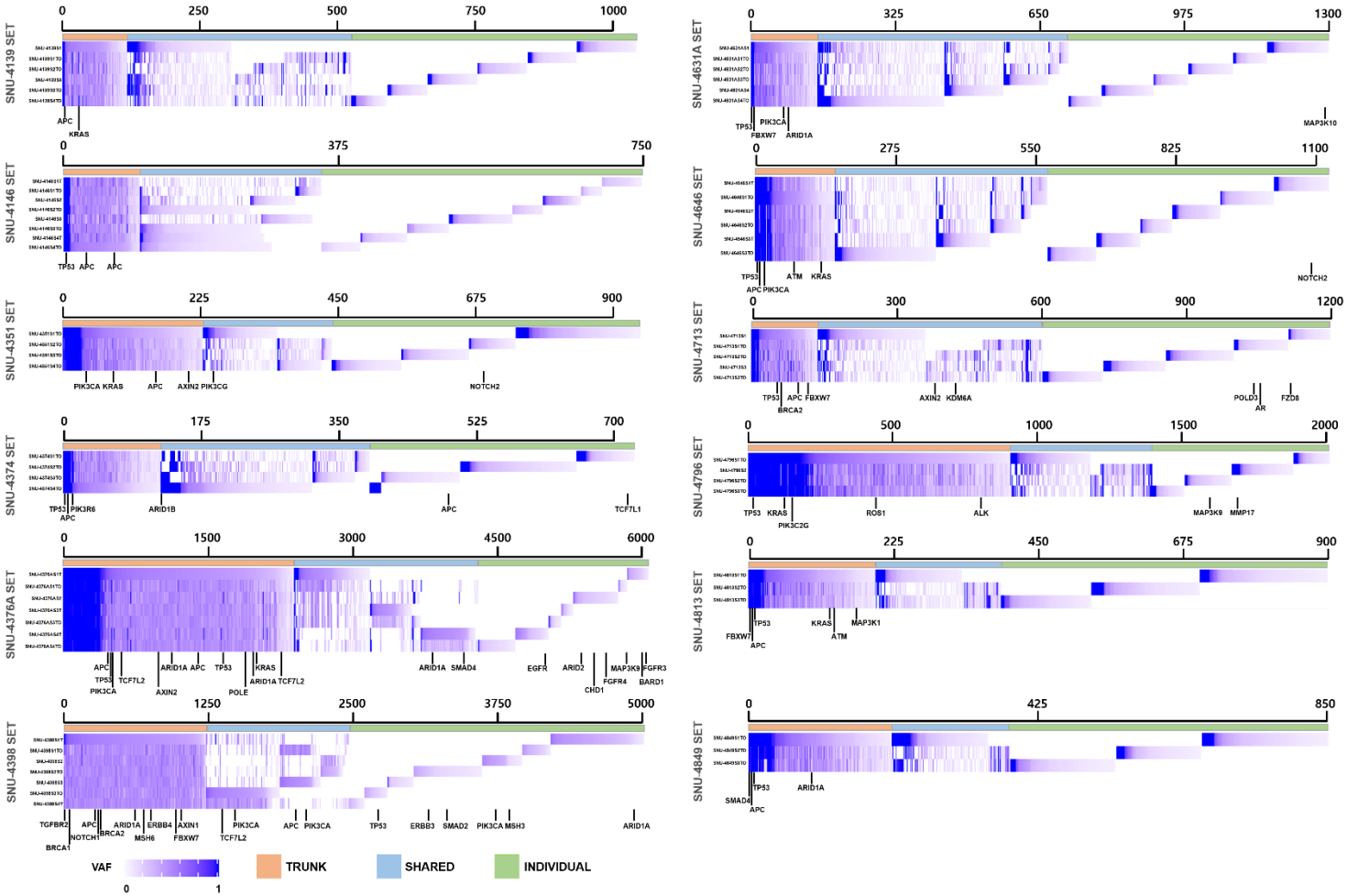

SNU-4139 SET
SNU-4146 SET
SNU-4351 SET
SNU-4374 SET
SNU-4376A SET
SNU-4398 SET
SNU-4631A SET
SNU-4646 SET
SNU-4713 SET
SNU-4796 SET
SNU-4813 SET
SNU-4849 SET
VAF
TRUNK
SHARED
INDIVIDUAL

Figure 12. Multiregion mutation profiles of twelve CRC series. Variant allele frequencies of all alternations of twelve CRCs were demonstrated with a heatmap for each case. Three different classes of mutations (trunk, shared, and individual) were indicated by top colored bars. Representative mutations that were reported as pathogenic were designated under each heatmap.

Gene Expression Analysis

We have normalized our gene expression data with normal mucosa sample. Since cell lines and organoid cultures comprise mainly epithelial cells, we have sorted our expression data with known epithelial cell markers (43, 44) to prevent potential bias from mesenchyme, blood vessels and immune cells surrounding normal tissues. Figure 13 shows the correlation heatmap of the organoid samples. Normal mucosa samples clustered together, while the tumor-derived organoids presented more heterogeneity. While samples were readily clustered in accordance with their tissue origin, several 2D cell lines (SNU-4796S1~S3) were located irrespective of their parental source, suggesting the culture method might affect RNA expressions of specific samples. We also screened differentially expressed genes between normal tissues and tumor organoids. Gene enriched in tumor organoids involved cancer-related genes such as AMHR2(45), PRDM2 (46) and VTI1A (47). Down-regulated genes in tumor organoids likewise included CRC-associated genes such as CLDN1(48, 49), DPEP1 (50, 51) and CXCL3 (52, 53) (Figure 14). We also screened RNA expression of MMR genes (MLH1, MLH3,

MSH2, MSH3, MSH6, PMS1 and PMS2). MLH1 expression was completely absent from SNU-4398 series and significantly decreased in SNU-4796 series (Figure 15). Neither of series had pathogenic mutation in MLH1, which suggested MLH1 expression was affected by methylation status.

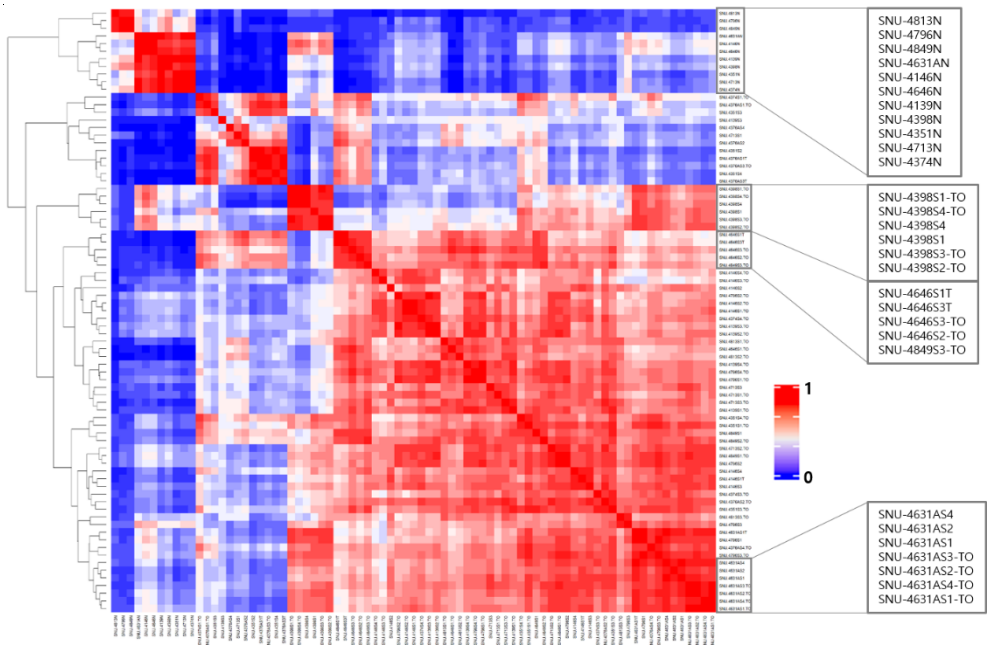


Figure 13. Distance plot of tumor organoids compared to normal tissue based on the top 10% of genes with respect to standard deviation (516 genes). Normal mucosa samples clustered together, while the tumor-derived organoids presented more heterogeneity. Colors symbolize pairwise Pearson correlations once the mRNA expression values were logged for every gene. The hierarchical clustering is based on one minus correlation distance. The affix N = normal, T = tumor.

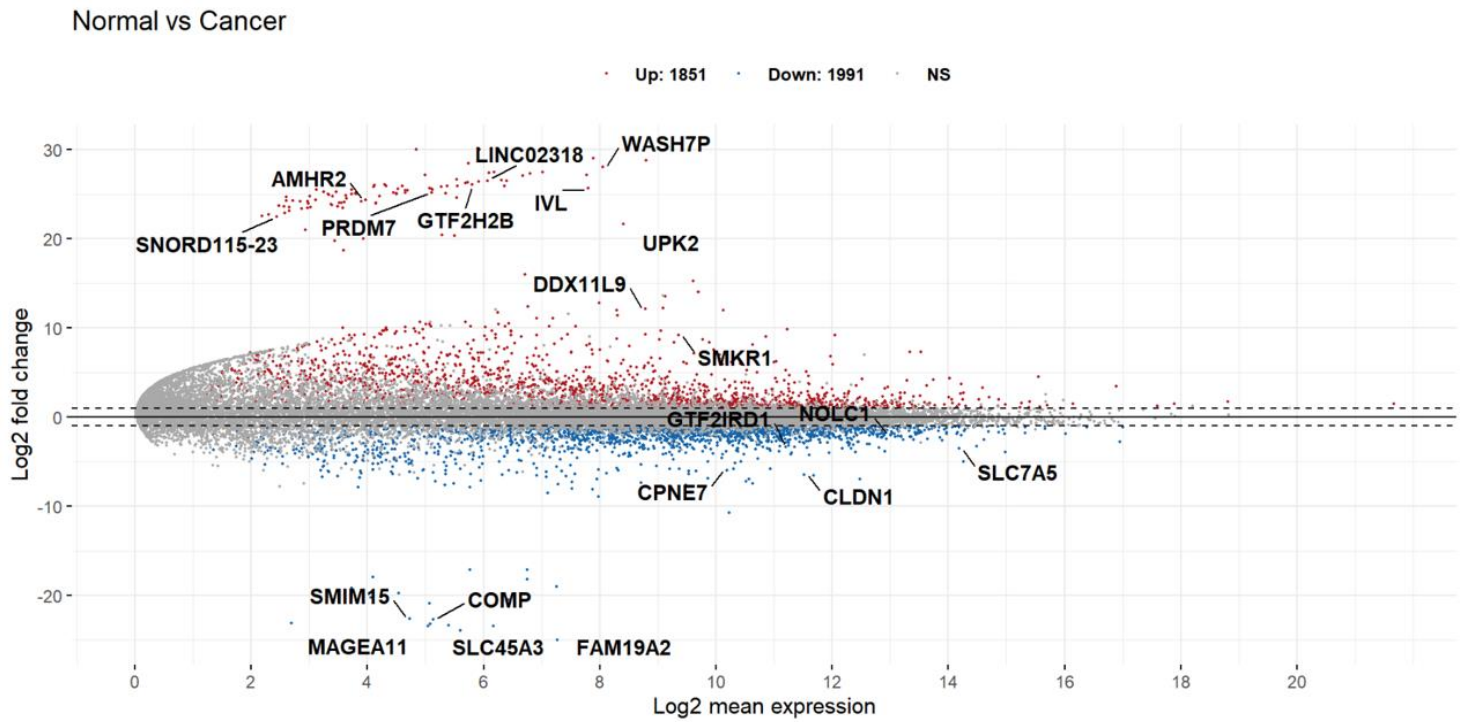


Figure 14. MA plot of logged normal versus tumor gene expression. P−values are computed with the R package limma, by comparing normal versus tumor gene expression. Cancer associated genes, e.g. AMHR2, PRDM2, VTI1A, CLDN1, DPEP1, CXCL3 are shown in the top half.

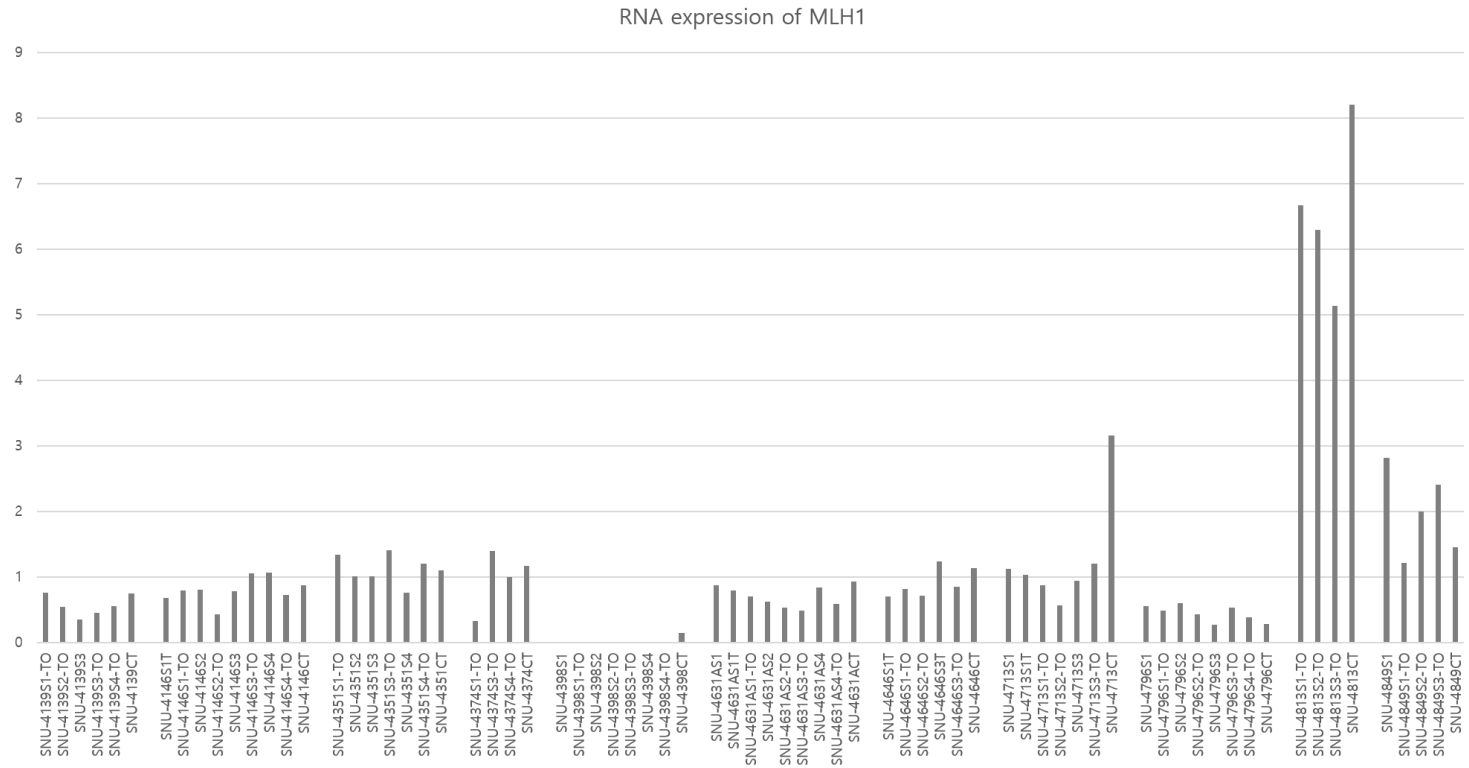


Figure 15. Boxplot of relative RNA expression of MLH1 gene. MLH1 expression was completely absent from SNU-4398 series and significantly decreased in SNU-4796 series.

Several CRC classifications have been proposed based on gene expression profiling (25, 26, 54). Recently, it has been reported that CMS classification is applicable to preclinical models such as cell lines and PDX models (27). We applied our RNA expression data from cell lines and organoid samples to CMS subtyping using an R package, CMScaller developed by Sveen et al (27). We used raw readcounts as a direct input with 'RNA-seq=True' setting in the program. A total of 82 RNA expression data from original tumor tissues, cell lines and organoids was subjected to subtyping (Figure 16, Table 7). Samples were distributed across the subtypes, with the CMS type 2 (n = 33) being most commonly represented. Seven samples were not affiliated to any of the CMS subtypes. The expression levels of gene sets that determine the CMS type were analyzed to confirm the specific pathways that are known to up- or down-regulated in each CMS types. For instance, the samples that were categorize as CMS4 displayed significantly increased EMT and TGF-β pathway related gene expression whereas gene sets that were involved in differentiation had diminished expression (Figure 17, Table 8)

There existed a significant intra-heterogeneity within a sample set (Figure 18, Table 7). For instance, the cancer tissue of patient SNU-4351 was classified as CMS type 3 while its derivate displayed CMS1 (SNU-4351S2), CMS3 (SNU-4351S1-TO, SNU-4351S3-TO, SNU-4351S4-TO), and CMS4 (SNU-4351S3, SNU-4351S4). In SNU-4351 case, only organoid samples retained the original tumor subtype. On the contrary, only cell line samples maintained the subtype of original tissue in SNU-4713 case.

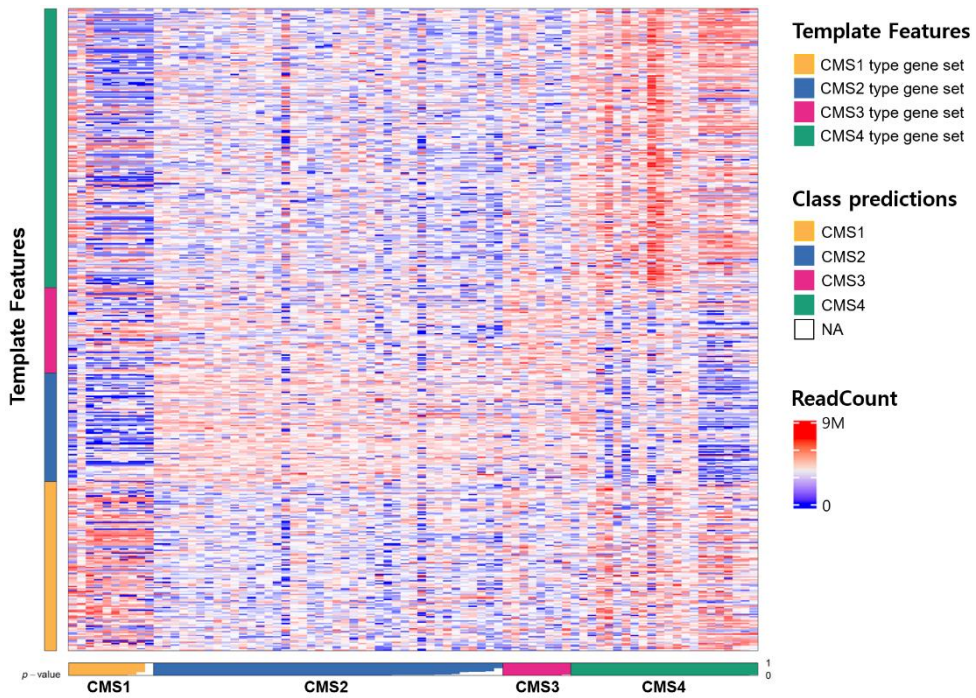


Figure 16. CMS classification of twelve CRC series. A total of 82 RNA expression data from original tumor tissues, cell lines and organoids was subjected to subtyping. Samples were distributed across the subtypes, with the CMS type 2 (n = 33) being most commonly represented. Seven samples were not affiliated to any of the CMS subtypes. Within each subtype, samples are sorted by their mean gene expression for the signature genes associated with that specific subtype.

Table 7. CMS classification of twelve CRC series.

Name	prediction	d.CMS1	d.CMS2	d.CMS3	d.CMS4	p.value	FDR
SNU.4139_CT	CMS4	0.788451997	0.627204597	0.66533751	0.598246362	0.001	0.001883721
SNU.4139S1.TO	CMS2	0.786697233	0.658894177	0.707636203	0.706607634	0.025974026	0.032367632
SNU.4139S2.TO	CMS2	0.754206366	0.632660072	0.668104991	0.708442552	0.001	0.001883721
SNU.4139S3	CMS4	0.68467436	0.624848381	0.635277834	0.578138174	0.001	0.001883721
SNU.4139S3.TO	CMS2	0.764767783	0.63287946	0.665946097	0.676113139	0.001	0.001883721
SNU.4139S4.TO	CMS2	0.767146049	0.647324021	0.689327243	0.709654603	0.001	0.001883721
SNU.4146_CT	CMS4	0.716944476	0.619814721	0.65441462	0.607209895	0.001	0.001883721
SNU.4146S1.TO	CMS2	0.772900582	0.632928515	0.673621244	0.783655944	0.002997003	0.004668409
SNU.4146S1T	NA	0.79643347	0.683119185	0.708674892	0.83210646	0.321678322	0.329822077
SNU.4146S2	CMS2	0.789700016	0.658519291	0.691147425	0.804665254	0.002997003	0.004668409
SNU.4146S2.TO	CMS2	0.727289575	0.616138406	0.695578683	0.758466483	0.001	0.001883721
SNU.4146S3	NA	0.724866981	0.68696434	0.714402755	0.795685189	0.282717283	0.297403895
SNU.4146S3.TO	CMS2	0.69147317	0.620447009	0.698597476	0.724738006	0.001	0.001883721
SNU.4146S4	NA	0.742147106	0.682334196	0.725801134	0.8057466	0.254745255	0.275124875
SNU.4146S4.TO	CMS2	0.757171576	0.64152176	0.712495517	0.777757506	0.001	0.001883721
SNU.4351_CT	CMS3	0.67084261	0.706982755	0.632774499	0.662692691	0.001	0.001883721
SNU.4351S1.TO	CMS3	0.735328683	0.71261332	0.656111182	0.805000496	0.004995005	0.00735628
SNU.4351S2	CMS1	0.630604809	0.792896594	0.757529993	0.632978925	0.001	0.001883721
SNU.4351S3	CMS4	0.586034704	0.748807707	0.659039805	0.565253207	0.001	0.001883721
SNU.4351S3.TO	CMS3	0.722351847	0.705787153	0.63163964	0.787143162	0.001998002	0.00330282
SNU.4351S4	CMS4	0.608394647	0.787385673	0.749418107	0.596007073	0.001	0.001883721
SNU.4351S4.TO	CMS3	0.631227217	0.735102963	0.630855987	0.73363444	0.001	0.001883721
SNU.4374_CT	CMS4	0.741508827	0.649435568	0.707005333	0.61878807	0.001	0.001883721
SNU.4374S1.TO	CMS4	0.688381212	0.813891238	0.746334811	0.641475891	0.001	0.001883721
SNU.4374S3.TO	CMS2	0.819404413	0.665757708	0.731870626	0.804645875	0.016983017	0.021835308
SNU.4374S4.TO	CMS2	0.790219914	0.605092084	0.684075453	0.751493608	0.001	0.001883721
SNU.4376AS1.TO	CMS4	0.65966489	0.809245009	0.734231765	0.604113584	0.001	0.001883721
SNU.4376AS1T	CMS4	0.633756831	0.809193821	0.748848647	0.601172898	0.001	0.001883721
SNU.4376AS2	CMS4	0.659109933	0.791745387	0.795249007	0.595620612	0.001	0.001883721
SNU.4376AS2.TO	CMS2	0.797264153	0.61212764	0.710224891	0.783521728	0.001	0.001883721
SNU.4376AS3.TO	CMS4	0.668084905	0.803173817	0.743509399	0.63401979	0.001	0.001883721
SNU.4376AS3T	CMS4	0.636906038	0.806633683	0.757107915	0.609479563	0.001	0.001883721
SNU.4376AS4	CMS4	0.635256529	0.782223614	0.746707048	0.575117114	0.001	0.001883721
SNU.4376AS4.TO	CMS4	0.686817512	0.779330918	0.696651191	0.66425416	0.033966034	0.041063414

Continued

Name	prediction	d.CMS1	d.CMS2	d.CMS3	d.CMS4	p.value	FDR
SNU.4398_CT	CMS1	0.639816	0.815485	0.733053	0.709703	0.001998	0.003303
SNU.4398S1	CMS1	0.656758	0.77288	0.725171	0.813407	0.005994	0.00867
SNU.4398S1.TO	NA	0.673374	0.791669	0.718868	0.815424	0.268731	0.286411
SNU.4398S2	CMS1	0.64474	0.793366	0.718778	0.813361	0.001998	0.003303
SNU.4398S2.TO	CMS1	0.647692	0.778048	0.717118	0.791133	0.015984	0.020882
SNU.4398S3.TO	CMS1	0.648229	0.780154	0.734574	0.808278	0.003996	0.006107
SNU.4398S4	NA	0.71898	0.778864	0.724422	0.838123	0.998002	0.998002
SNU.4398S4.TO	CMS1	0.664156	0.778223	0.723785	0.803108	0.050949	0.058955
SNU.4631A_CT	CMS3	0.715983	0.687946	0.625402	0.667377	0.001	0.001884
SNU.4631AS1	NA	0.751348	0.676775	0.752268	0.726494	0.287712	0.298778
SNU.4631AS1.TO	CMS2	0.681698	0.660153	0.68403	0.779586	0.026973	0.033103
SNU.4631AS1T	CMS2	0.719812	0.671011	0.692673	0.803735	0.063936	0.072941
SNU.4631AS2	NA	0.783832	0.685905	0.764247	0.833027	0.559441	0.566434
SNU.4631AS2.TO	CMS3	0.6818	0.690334	0.66272	0.778044	0.076923	0.085353
SNU.4631AS3.TO	CMS2	0.685517	0.66337	0.697298	0.748939	0.06993	0.078671
SNU.4631AS4	CMS2	0.753652	0.670186	0.716065	0.791886	0.103896	0.113724
SNU.4631AS4.TO	CMS2	0.721927	0.664214	0.685093	0.783088	0.038961	0.045737
SNU.4646_CT	CMS4	0.711995	0.69151	0.686344	0.547748	0.001	0.001884
SNU.4646S1.TO	CMS2	0.739264	0.650818	0.706818	0.668763	0.006993	0.009766
SNU.4646S1T	CMS4	0.614948	0.633277	0.668605	0.578115	0.001	0.001884
SNU.4646S2.TO	CMS2	0.693575	0.658425	0.682931	0.68121	0.00999	0.013265
SNU.4646S3.TO	CMS2	0.708594	0.640224	0.663842	0.685198	0.001	0.001884
SNU.4646S3T	CMS1	0.62942	0.662223	0.691355	0.645609	0.001	0.001884
SNU.4713_CT	CMS4	0.72067	0.758854	0.751973	0.464073	0.001	0.001884
SNU.4713S1	CMS4	0.620216	0.726951	0.713845	0.410327	0.001	0.001884
SNU.4713S1.TO	CMS2	0.735663	0.629875	0.686893	0.745124	0.001	0.001884
SNU.4713S1T	CMS4	0.646381	0.64514	0.681312	0.493305	0.001	0.001884
SNU.4713S2.TO	CMS2	0.730805	0.646788	0.67352	0.785397	0.002997	0.004668
SNU.4713S3	CMS3	0.742627	0.666028	0.635594	0.641641	0.001	0.001884
SNU.4713S3.TO	CMS2	0.73044	0.657902	0.693208	0.709639	0.025974	0.032368
SNU.4796_CT	CMS2	0.780818	0.703493	0.733022	0.722336	0.001998	0.003303
SNU.4796S1	CMS2	0.749091	0.653368	0.686968	0.77343	0.006993	0.009766
SNU.4796S1.TO	CMS2	0.737654	0.639037	0.698621	0.66505	0.001	0.001884
SNU.4796S2	CMS2	0.730207	0.634037	0.683628	0.720419	0.001	0.001884
SNU.4796S2.TO	CMS2	0.793223	0.623233	0.69899	0.716254	0.001	0.001884
SNU.4796S3	CMS3	0.784824	0.664982	0.639146	0.799502	0.001	0.001884
SNU.4796S3.TO	CMS2	0.766645	0.626905	0.665592	0.753542	0.001	0.001884
SNU.4796S4.TO	CMS2	0.773209	0.653111	0.744101	0.713892	0.00999	0.013265
SNU.4813_CT	CMS2	0.767759	0.668242	0.746605	0.70959	0.00999	0.013265
SNU.4813S1.TO	CMS2	0.666415	0.638671	0.714567	0.665598	0.001998	0.003303
SNU.4813S2.TO	CMS4	0.613592	0.679759	0.712868	0.603963	0.001	0.001884
SNU.4813S3.TO	CMS2	0.688622	0.654301	0.721225	0.717321	0.001998	0.003303
SNU.4849_CT	CMS2	0.776681	0.713321	0.760105	0.76406	0.035964	0.04284
SNU.4849S1	CMS4	0.649122	0.619165	0.66972	0.516448	0.001	0.001884
SNU.4849S1.TO	CMS2	0.725223	0.63988	0.653182	0.691632	0.001	0.001884
SNU.4849S2.TO	CMS2	0.697064	0.65767	0.658705	0.665373	0.004995	0.007356
SNU.4849S3.TO	CMS4	0.700577	0.636565	0.668868	0.633529	0.001	0.001884

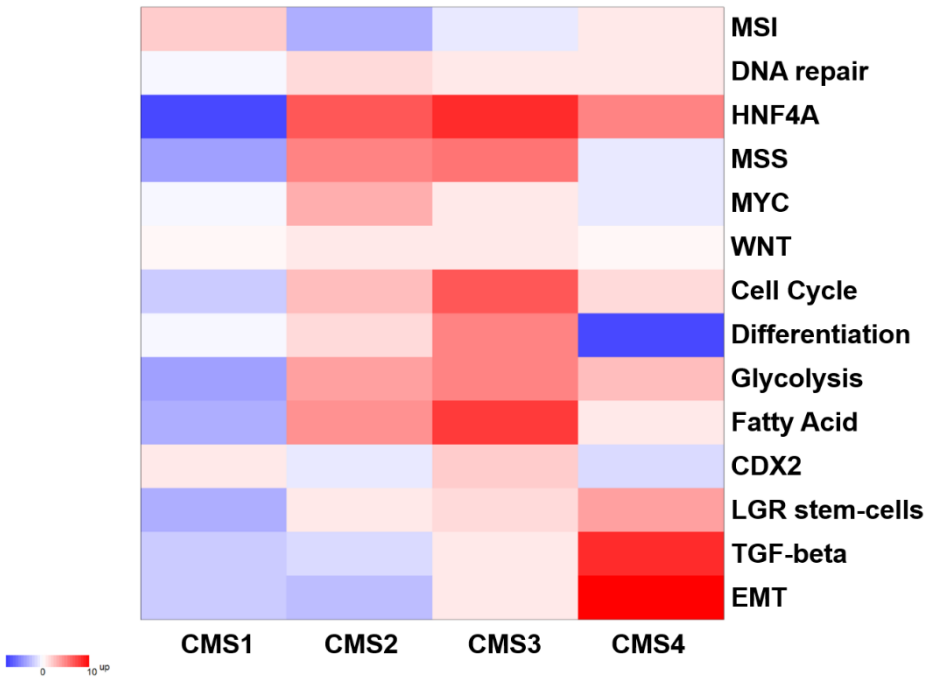


Figure 17. Gene set enrichment analysis of CMS classification. The expression levels of gene sets that determine the CMS type were analyzed to confirm the specific pathways that are known to up- or down-regulated in each CMS types. For instance, the samples that were categorize as CMS4 displayed significantly increased EMT and TGF-β pathway related gene expression whereas gene sets that were involved in differentiation had diminished expression.

Table 8. Gene set enrichment analysis of CMS classification

	CMS1. NGenes	CMS1. Direction	CMS1. PValue	CMS1. FDR	CMS2. NGenes	CMS2. Direction	CMS2. PValue	CMS2. FDR
EMT	199	Down	0.0243	0.0378	199	Down	0.0046	0.0080
TGF-Beta	60	Down	0.0177	0.0309	60	Down	0.0625	0.0875
LGR5 stem-cells	62	Down	0.0032	0.0089	62	Up	0.2266	0.2643
CDX2	36	Up	0.7447	0.9603	36	Down	0.4806	0.5074
fatty acids	158	Down	0.0022	0.0076	158	Up	0.0001	0.0004
glycolysis	200	Down	0.0010	0.0045	200	Up	0.0009	0.0033
differentiation	628	Down	0.9082	0.9603	628	Up	0.0795	0.1012
cell cycle	200	Down	0.0169	0.0309	200	Up	0.0038	0.0076
WNT	13	Up	0.9603	0.9603	13	Up	0.5074	0.5074
MYC	58	Down	0.9456	0.9603	58	Up	0.0028	0.0066
MSS	81	Down	0.0009	0.0045	81	Up	0.0000	0.0002
HNF4A	58	Down	0.0000	0.0000	58	Up	0.0000	0.0000
DNA repair	150	Down	0.8369	0.9603	150	Up	0.0615	0.0875
MSI	29	Up	0.0157	0.0309	29	Down	0.0025	0.0066

Continued

	CMS3. NGenes	CMS3. Direction	CMS3. PValue	CMS3. FDR	CMS4. NGenes	CMS4. Direction	CMS4. PValue	CMS4. FDR
EMT	199	Up	6E-01	7E-01	199	Up	9E-11	1E-09
TGF-Beta	60	Up	4E-01	5E-01	60	Up	2E-08	1E-07
LGR5 stem-cells	62	Up	9E-02	2E-01	62	Up	9E-04	2E-03
CDX2	36	Up	4E-02	7E-02	36	Down	1E-01	2E-01
fatty acids	158	Up	4E-08	3E-07	158	Up	4E-01	5E-01
glycolysis	200	Up	2E-05	6E-05	200	Up	1E-02	2E-02
differentiation	628	Up	6E-05	1E-04	628	Down	9E-08	4E-07
cell cycle	200	Up	4E-07	2E-06	200	Up	9E-02	2E-01
WNT	13	Up	7E-01	7E-01	13	Up	9E-01	9E-01
MYC	58	Up	4E-01	5E-01	58	Down	6E-01	8E-01
MSS	81	Up	5E-06	2E-05	81	Down	7E-01	8E-01
HNF4A	58	Up	8E-09	1E-07	58	Up	6E-05	2E-04
DNA repair	150	Up	3E-01	4E-01	150	Up	3E-01	4E-01
MSI	29	Down	3E-01	5E-01	29	Up	6E-01	7E-01

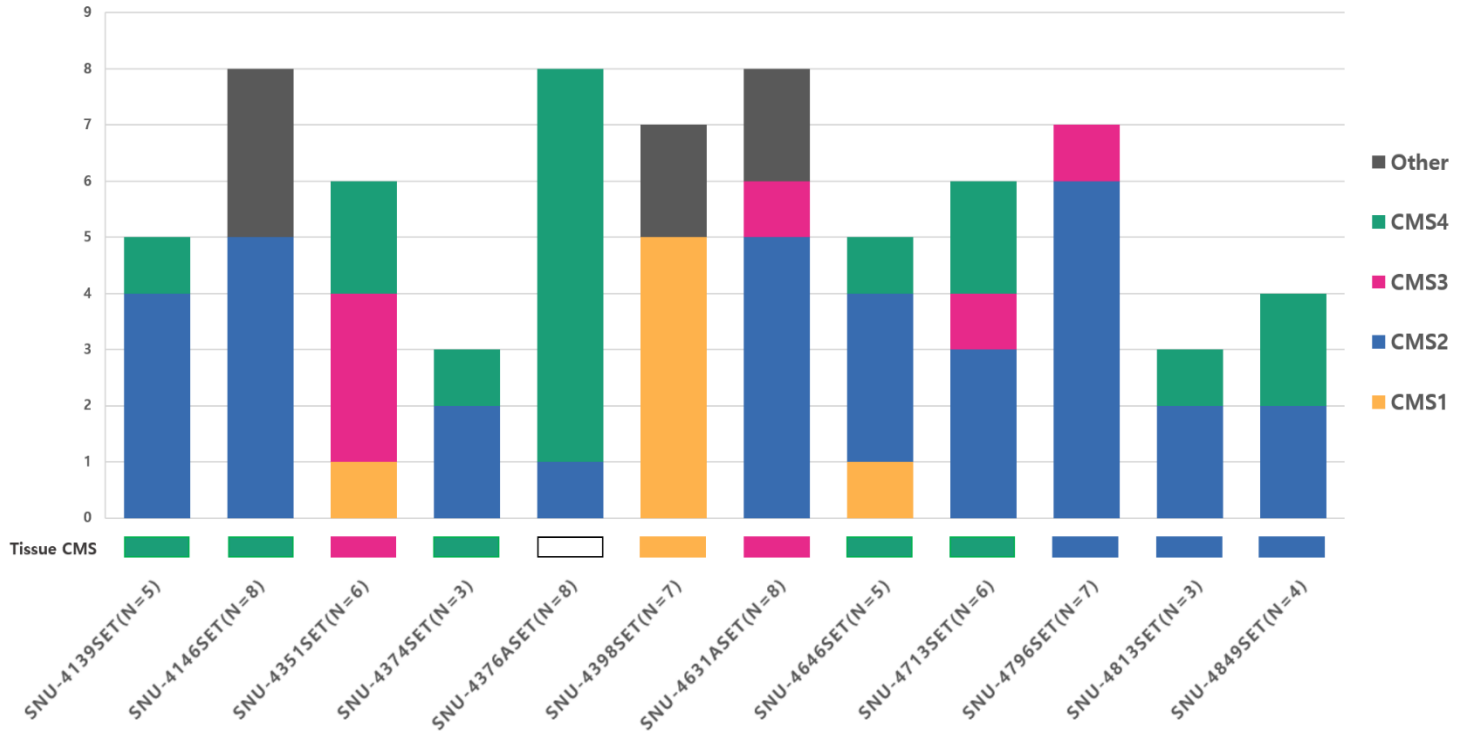


Figure 18. Presence of intra-heterogeneity within a sample set

Patient Derived Organoids Enables Drug Response Prediction

A 24 compound library was assembled for screening, including anti-metabolites (n = 3), kinase inhibitors (n = 7), histone deacetylase inhibitors (n = 2), alkylating inhibitors (n = 1), topoisomerase inhibitors (n = 1), growth factor receptor inhibitors (n = 2), natural compounds (n = 4), and miscellaneous (n = 4). In total, 40 of 42 tumor organoids and 20 of 23 cell lines from 12 patients were successfully screened in experimental duplicate, generating >1200 measurements of organoid-drug and cell line-drug interactions.

As a first validation, the grouping of compounds based on their AUC values confirmed a various range of responses across the cell lines and organoids, and identified 5 major sub-groups in accordance with compounds (Figure 19). One group was related with relatively high sensitivity to EGFR/RAS/RAF/MEK/ERK pathway targeting drugs with HDAC targeting drugs, in contrast to the group exhibiting intermediate sensitivity. Other group displayed intermediate sensitivity to anti-metabolites and topoisomerase inhibitors, in contrast to the cluster exhibiting insensitivity. The final group

involved phytochemicals which had insensitivity across the entire samples (Figure 19).

The multifocal samples tended to cluster together according to their tumor origin with a variation in several compounds. Even within a same cluster, cell lines were grouped adjacent to each other, suggesting that the culture method impacted the drug sensitivities. We identified clustering of drugs that inhibit the PI3K/MTOR and MEK signaling pathways, and compounds with analogous molecular targets had corresponding responses across the cell lines and organoids. For example, a comparable sensitivity pattern was perceived for the PI3K inhibitors Buparlisib and Apitolisib, the MEK inhibitors Trametinib, and HDAC inhibitors belinostat (Figure 19).

We also observed few compounds with diverse sensitivities regardless of an obvious genetic biomarker. For instance, a group of organoids was exceedingly sensitive to the PI3K inhibitor Buparlisib. Likewise, we identified different subsets of organoids which are particularly sensitive to the MEK inhibitor, Trametinib and the HDAC inhibitor, Belinostat (Figure 19). To sum up, the utility and practicability of the in vitro cancer models for examining the

molecular bases of drug sensitivity were validated with the efficacious application of both cell lines and organoids in a methodical and unbiased high-throughput drug screening to determine clinically applicable biomarkers.

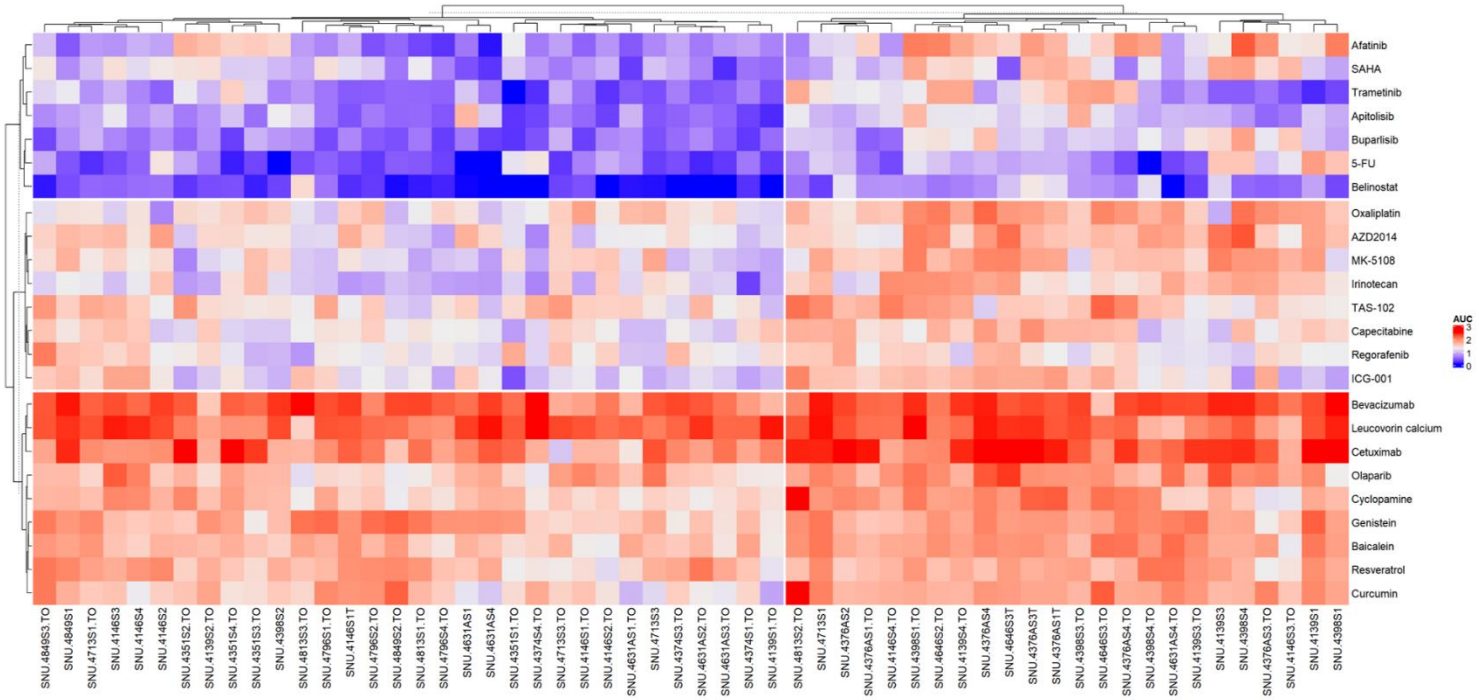


Figure 19. Heatmap of AUCs of all 24 compounds against 56 CRC derivate. Cell lines/Organoids have been clustered based on their AUCs values across the drug panel. The drug names are provided in the right panel.

In order to integrate genetic and transcriptomic analysis to drug response, we performed a multivariate analysis of variance (MANOVA) incorporating various elements such as mutational profiles, shapes of phylogenetic trees, CMS subtypes and culture methods to AUC values. Comprehensive datasets comprising essential factors were available for 56 derivate and used for this analysis. The MANOVA identified a subset of CMS types–drug associations as statistically significant. Consistent with epithelial and metabolic feature in CMS3, three anti–metabolite drugs (Capecitabine, 5–FU and TAS–102) displayed better response in CMS type 3 samples. Notably, oxaliplatin, one of traditional anti–metabolic drug did not show improved response in CMS3 samples (Figure 20).

As previously reported (55, 56), *KRAS* mutant samples showed resistance to the anti–EGFR inhibitors, cetuximab and afatinib. Interestingly, both cell lines and organoids that were derived from patient SNU–4398 displayed significant resistance to afatinib despite of wild type *KRAS*. We further inspected mechanisms beyond mutated *KRAS/NRAS/BRAF* in afatinib resistance. We detected

unique *BRAF_GTF2IRD1* fusion in SNU-4398 series. The fusion gene exclusively presented at tumor tissue and its derivate (Figure 22b). The naïve mRNA expression of both *BRAF* and *GTF2IRD1* presented at the normal tissue, which indicated that the fusion gene was somatic event. The mRNA expression of *BRAF* was increased in tumor tissue and its derivate (Figure 22b). The *BRAF_GTF2IRD1* fusion occurred only within SNU-4398 series among the 12 CRC series (Figure 22c). Similar fusion event, *GTF2I-BRAF* was reported previously to activate MAPK pathway in pilocytic astrocytoma (57), which may suggest genomic loci spanning BRAF (7q34) and GTF2IRD1 (7q11.23) are susceptible to chromosomal instability. Genomic and transcriptomic landscape of SNU-4398 series indicated that sub-clones of SNU-4398 series acquired *BRAF_GTF2IRD1* fusion gene (Figure 23). The 66,508,972bp deletion at chromosome 7 didn't affect the copy numbers between fusion junctions. In contrast, mRNA expression of BRAF was significantly increased, which might imply somatic fusion event occurred and formed the sub-clone with activated MAPK pathway and resistance to afatinib accordingly.

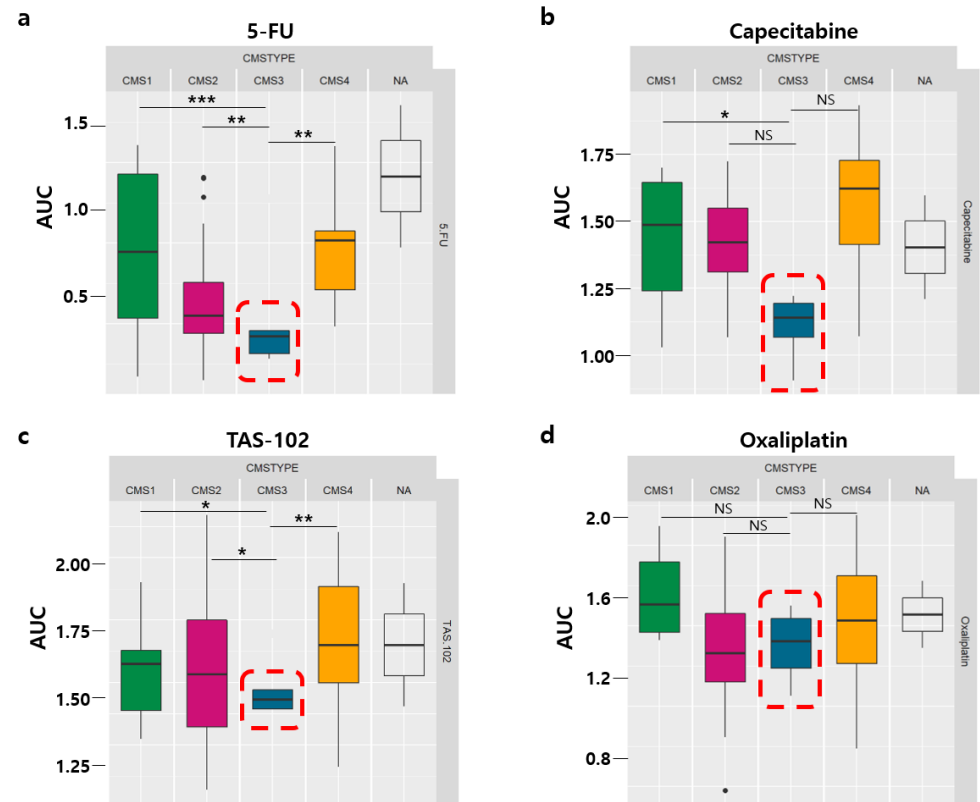


Figure 20. Drug sensitivity of CMS type 3 cancer to anti-metabolite drugs. Anti-metabolite drugs (a. 5-FU, b. Capecitabine, c. TAS-102 and d. Oxaliplatin) displayed better response in CMS type 3 samples except for Oxaliplatin.

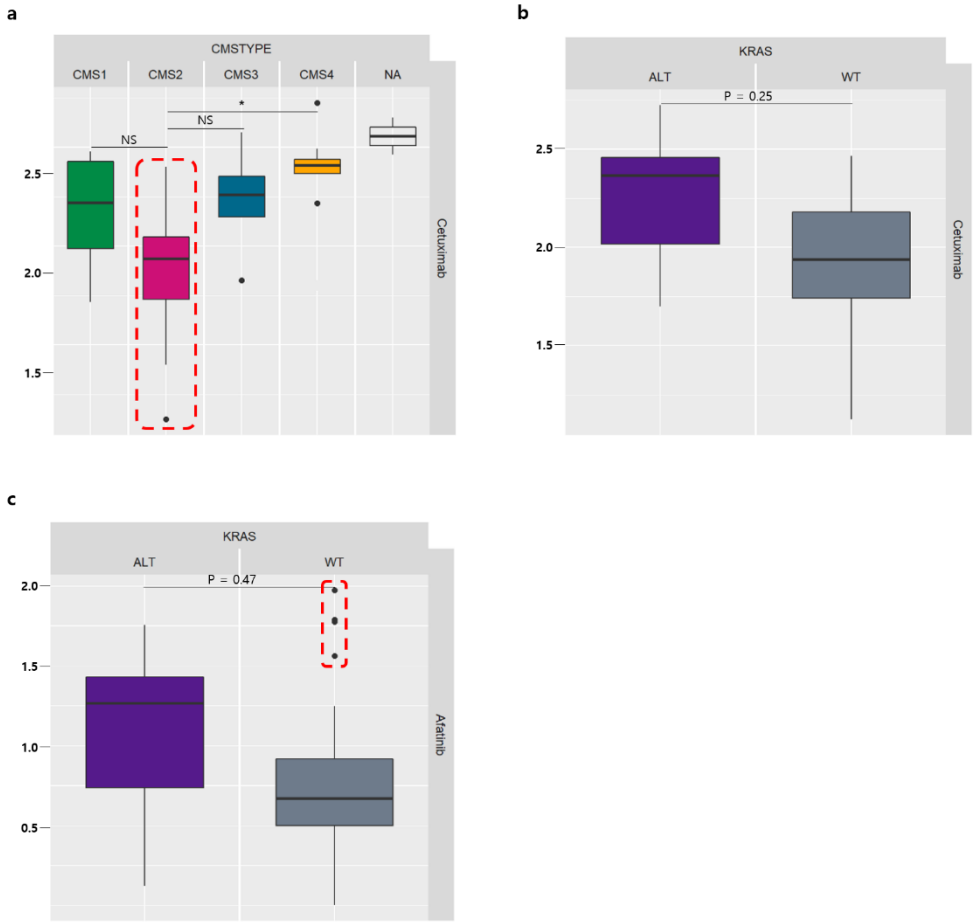


Figure 21. Drug sensitivity of CMS type 2 cancer to EGFR targeting drugs (a. CMS-Cetuximab, b. Kras-Cetuximab, and c. Kras-Afatinib). CMS type 2 cancer displayed a better response to Cetuximab compared to other types.

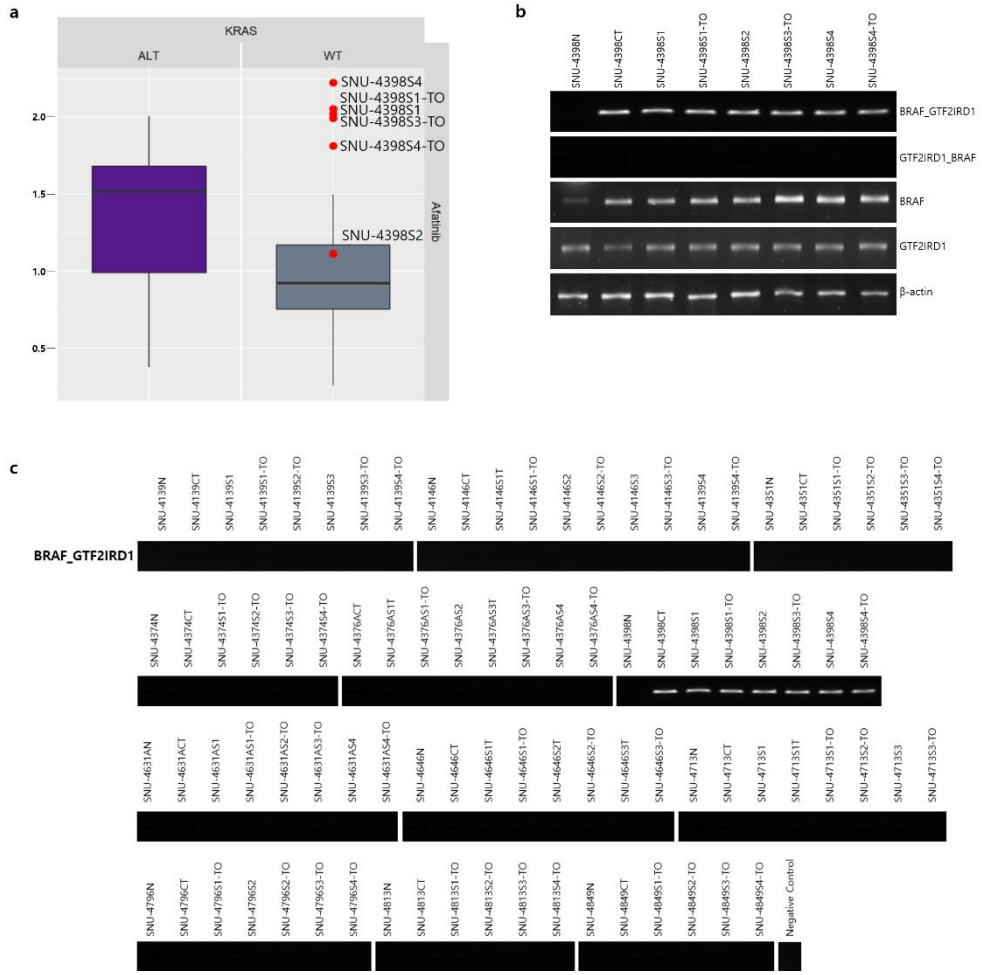


Figure 22. a. There existed resistance to afatinib within SNU-4398 series. b. Unique *BRAF_GTF2IRD1* fusion in SNU-4398 series presented at tumor tissue and its derivate. The naïve mRNA expression of both *BRAF* and *GTF2IRD1* presented at the normal tissue. c. The *BRAF_GTF2IRD1* fusion occurred only within SNU-4398 series among the 12 CRC series.

Table 9. SNU-4398 series harbored *BRAF_GTF2IRD1* fusion gene

SNUCODE	5end_gene	3end_gene	5end_gene_chrom	5end_gene_pos	3end_gene_chrom	3end_gene_pos	positional difference
SNU-4398CT	BRAF	GTF2IRD1	7	140487348	7	74004180	-66483168
SNU-4398S1	BRAF	GTF2IRD1	7	140487348	7	74004180	-66483168
SNU-4398S1-TO	BRAF	GTF2IRD1	7	140487348	7	74004180	-66483168
SNU-4398S2	BRAF	GTF2IRD1	7	140487348	7	74004180	-66483168
SNU-4398S2-TO	BRAF	GTF2IRD1	7	140487348	7	74004180	-66483168
SNU-4398S3-TO	BRAF	GTF2IRD1	7	140487348	7	74004180	-66483168
SNU-4398S4	BRAF	GTF2IRD1	7	140487348	7	74004180	-66483168
SNU-4398S4-TO	BRAF	GTF2IRD1	7	140487348	7	74004180	-66483168

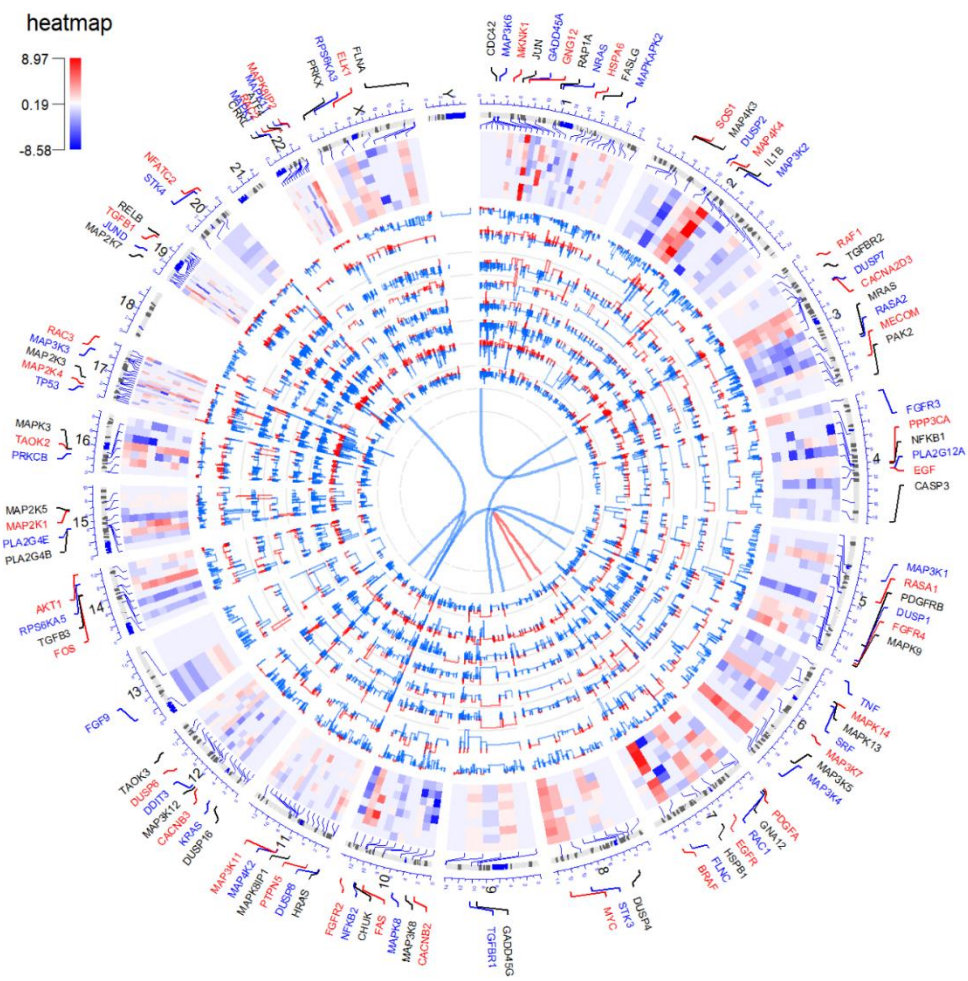

heatmap
8.97
0.19
-8.58
CDC42
MAP3K6
MKNK1
JUN
GADD45A
GNG12
RAP1A
NRAS
HSPA8
FASLG
MAPKAPK2
SOS1
MAP4K3
DUSP2
MAP4K4
IL1B
MAP3K2
RAF1
TGFBR2
DUSP7
CACNA2D3
MRAS
RASA2
MECOM
PAK2
FGFR3
PPP3CA
NFKB1
PLA2G12A
EGF
CASP3
MAP3K1
RASA1
PDGFRB
DUSP1
FGFR4
MAPK9
TNF
MAPK14
MAPK13
SRF
MAP3K7
MAP3K5
MAP3K4
PDGFA
GNA12
RAC1
EGFR
HSPB1
FLNC
BRAF
DUSP4
MYC
STK3
TGFBR1
GADD45G
FGFR2
NFKB2
CHUK
FAS
MAPK8
MAP3K8
CACNB2
MAP3K11
MAP4K2
MAP6IP1
PTPN5
DUSP6
HRAS
TAOK3
DUSP6
DDIT3
MAP3K12
CACNB3
KRAS
DUSP16
FGF9
AKT1
RPS6KA5
TGFB3
FOS
MAP2K5
MAP2K1
PLA2G4E
PLA2G4B
MAPK3
TAOK2
PRKCB
RAC3
MAP3K3
MAP2K3
MAP2K4
TP53
RELB
TGFB1
JUND
MAP2K7
NFATC2
STK4
MAPK6IP2
NFKBIA
MAPK1
CHUK
PRKX
RPS6KA3
ELK1
FLNA
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y

Figure 23. Genomic, transcriptomic landscape of SNU-4398 series. Labelled genes around chromosomes are involved in MAPK pathways. Heatmap indicates mRNA expressions of labelled genes. (From outer to inner circles represents SNU-4398-CT, SNU-4398-S1, SNU-4398S1-TO, SNU-4398-S2, SNU-4398-S2TO, SNU-4398-S3TO, SNU-4398S4, SNU-4398S4-TO respectively). Barplot indicates copy number alternations with same order with identical order with heatmap. The inner most linkage lines are fusion genes detected by both defuse and fusionCatcher programs. *BRAF_GTF2IRD1* fusion gene is highlighted with red color.

Discussion

Patient-derived tumor organoids have been widely used for personalized cancer medicine (58, 59), reflecting its value in both basic cancer research (60) and translational research (61). PDO cultures for in vitro modeling of original tumors have been applied to colorectal cancer as well (37). Colorectal cancers are assorted with heterogeneous sub-clones that were shaped by a Darwinian selection process (62). This molecular heterogeneity and phenotypic disparity construct a multifaceted clonal architecture, which supports significant features such as drug resistance and metastatic potential (63, 64).

Uncontrolled propagation of normal organoids within the tumor organoid population is one of major challenges in culturing tumor organoids (29). Outgrowth of normal organoid necessarily impedes the accurate computation of genomic treats of tumor organoids including the copy number variation and the somatic mutation detection. Therefore, we have confirmed that growth of normal organoid is arrested when tumor organoid culture medium is used for

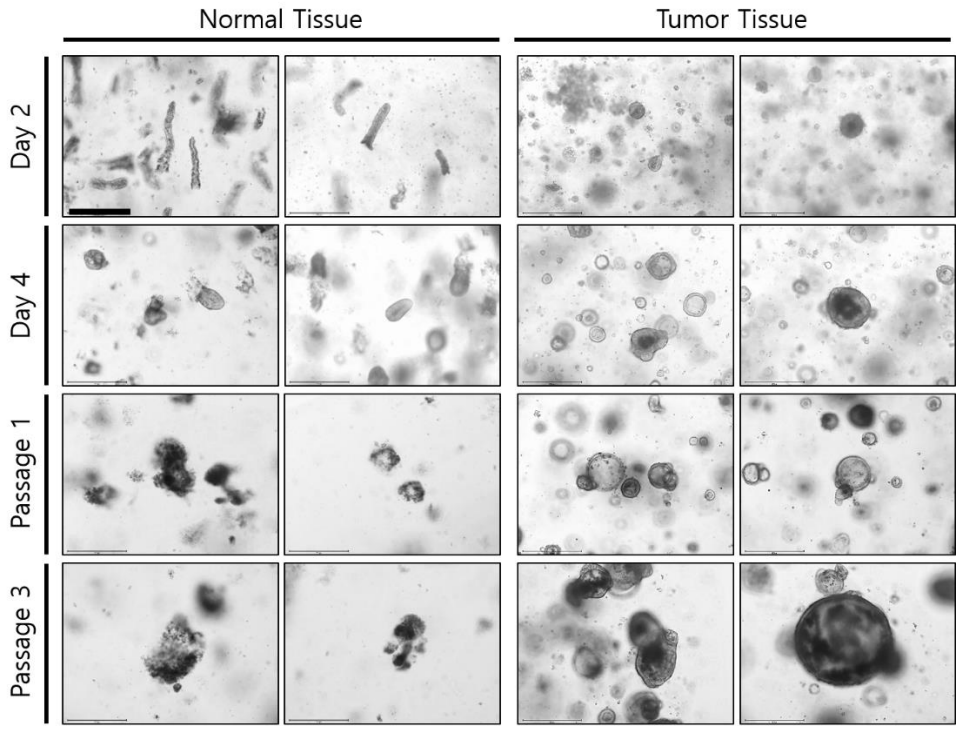
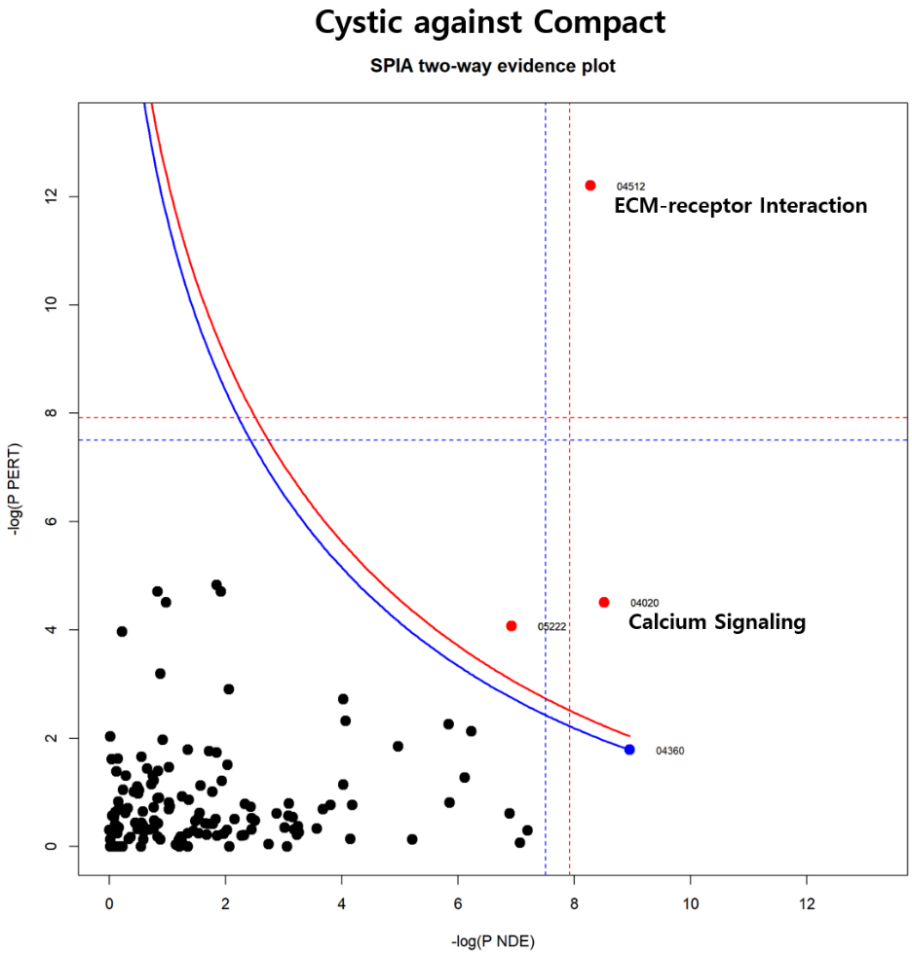


Figure 24. Cultivation of both primary normal (left) and the corresponding tumor tissue (right). The structure of normal crypts started to be dissociated within three days and cryptic architecture is completely destructed after passaging. Scale bars = 500 μM.

multiple tumor organoid cultures. In general, the configuration of normal crypts started to be dissociated within three days and cryptic architecture is completely destructed after passaging (Figure 24).

Established tumor organoids displayed morphological variations (Figure 1-3). Each organoid is categorized as cystic, compact and mixed population in accordance with the H&E staining and ICC results. In order to identify if the morphological differences were derived from transcriptomic variances, we performed SPIA analysis. SPIA evidence plot was drawn to cystic structure organoids against compact structure organoids (Figure 25). The mRNA expressions of cystic organoids displayed inhibited ECM-receptor interaction signaling and calcium signaling compared to solid organoids.



Name	ID	pSize	NDE	pNDE	Status
ECM-recept	4512	85	36	0.000254	Inhibited
Calcium si	4020	183	67	0.000202	Inhibited

Figure 25. SPIA two-way evidence plot of cystic against compact organoids. Each pathway in the database is represented by a single dot. The pathways at the right side of the red and blue curves are considered as significant after Bonferroni and FDR correction of the global p values respectively.

In this research, multiregional specimens from primary colorectal tumor tissues discovered a significant degree of ITH. Although genomic ITH has been investigated across multiple cancer types involving a study of colorectal adenomas (65), our report is the first study on the ITH of colorectal cancer genomes with a comprehensive interpretation on consensus molecular subtypes and evolutionary trajectories of primary lesions obtained from the same patient. Our variant allele frequency (VCF) analysis of sub-clonal configuration indicated that all tumor mass was heterogeneous at the time of surgical resection, which is mirrored to variable degree in the matched cell line and organoid cultures. Since comprehending the factors that affect tumor evolution and heterogeneity which influences on drug sensitivities is absolutely substantial to completely understand their potential for predicting patient responses, the capability of organoid technique to maintain sub-clonal features facilitate more precise modeling of patient responses to therapy.

Capecitabine (CAP) is an orally administered systemic prodrug of 5'-deoxy-5-fluorouridine (5'-DFUR) which is enzymatically

converted to 5-fluorouracil (5-FU). In the liver, CAP is hydrolyzed by carboxyesterase to 5'-deoxy-5-fluorocytidine (5'-DFCR). Cytidine deaminase, an enzyme found in most tissues, including tumors, subsequently converts 5'-DFCR to 5'-deoxy-5-fluorouridine (5'-DFUR). The enzyme, thymidine phosphorylase (dThdPase), then hydrolyzes 5'-DFUR to the active drug 5-FU (66). Although the presence of essential enzymes for the conversion of CAP to 5-FU within the tumor organoid medium is in question, our result indicated that the primary form of CAP exhibited relatively moderate cytotoxicity on tumor cell lines/organoids (median AUC = 1.42, range 0.91 to 1.93) compared to 5-FU (median AUC = 0.64, range 0.15 to 1.86). Other studies have also indicated the primary form of CAP had cytotoxicity and genotoxicity (67-69). Besides, in terms of CMS classifications, CMS type 3 cell lines/organoids exhibited better response to Capecitabine in parallel with 5-FU. Therefore, we determined to include the results of CAP in our drug response data.

Our cell line/organoid drug screening assay produced replicable drug sensitivity data, which reflected positive association of biological

replicates and reproducible sensitivity of compounds targeting the identical cellular molecules. By associating genetic/transcriptomic and drug sensitivity data, we were able to ratify the sensitivity of cetuximab in a subset of KRAS wild-type organoids reproducing observations from the clinic (70) as well as efficacy anti-metabolite compounds in CMS type 3 cell lines and organoids. We also detected that the activity of cetuximab in a subset of CMS type 2 cell lines and organoids, which was in line with previous report (27, 71). Interestingly, this tendency was diminished in a drug response in 2D cell lines, suggesting that organoid functions better as a preclinical model. We have integrated the clonal evolution event with drug response. We assumed that the adjacent sub-clones in the phylogenetic tree have comparable responses to certain drugs. Figure 26 supported this notion. Figure 26a indicated that SNU-4139S3-TO and SNU-4139S4TO sub-clones had close evolutional distance as well as drug responses. SNU-4139S1 and SNU-4139S3 sub-clones also presented analogous mutational profiles in parallel with drug responses. The similar pattern was observed in SNU-4631A series. Nevertheless, not every series have correlation

between clonal evolution and drug responses. In SNU-4374 series, sub-clones have identical distance in phylogenetic tree (Figure 11). In these cases, we performed western blotting to find if there are differences in protein level. Interestingly, the protein levels were significantly disparate in both SNU-4146 and SNU-4374 series (Figure 27a, b). In SNU-4146 series, AKT-mTOR pathway was specifically activated in organoid samples (Figure 27a). The response of AZD2014, the reagent inhibiting mTOR activation had good response to organoid groups. Similar pattern was observed in SNU-4374 series. AKT-mTOR pathway was explicitly up-regulated in SNU-4374S4-TO samples and AZD2014 showed better response (Figure 27b). Figure 27b suggested that heterogeneity existed not only between different culture types but also within a sample culture type. We also confirmed that targeting trunk mutation is effective to all sub-clones (Figure 28). Trametinib is MEK inhibitor and reported to be effective to KRAS mutant CRC. All sub-clones of SNU-4139 series harbored KRAS mutation, and showed good response to trametinib. In contrast, drugs targeting shared or individual mutations such as afatinib and buparlisib had varying

response (Figure 28a). SNU-4631A series harbored EGFR, PIK3CA, and RAF1 mutations and showed good response to afatinib, buparlisib and trametinib (Figure 28b).

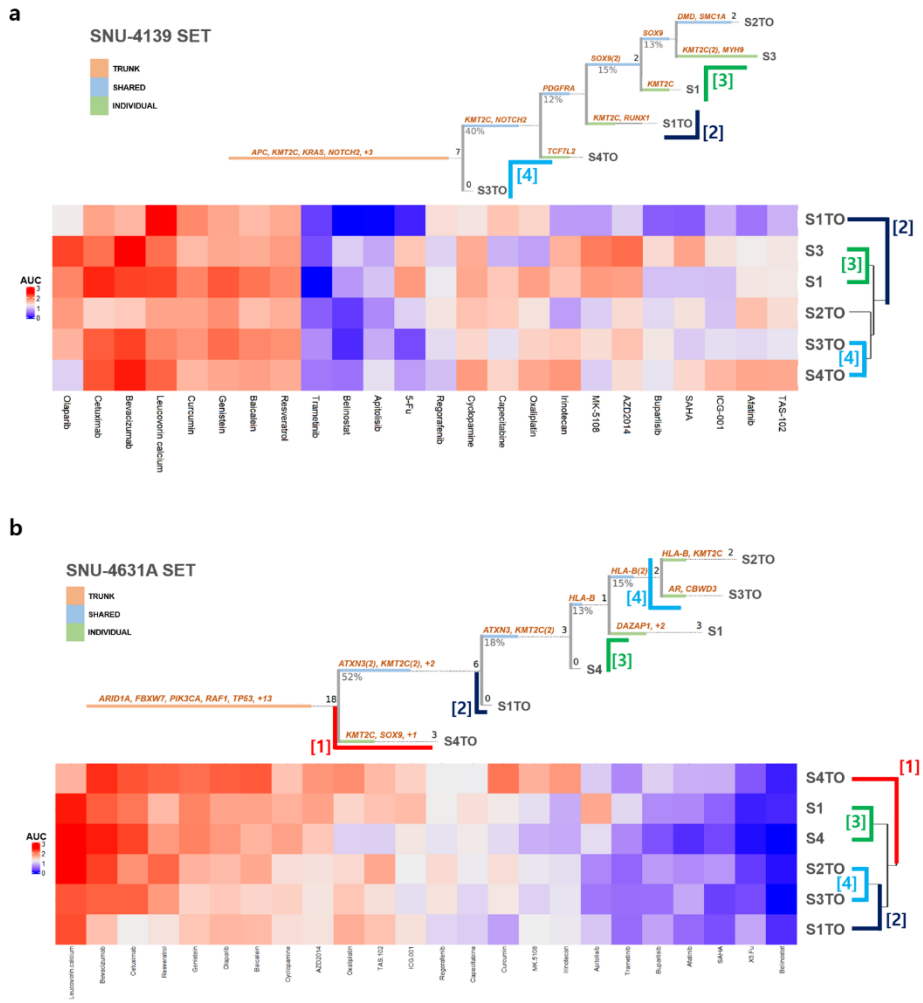


Figure 26. Clonal evolution may predict drug response. a. SNU-4139S3-TO and SNU-4139S4TO sub-clones had close evolutional distance as well as drug responses. SNU-4139S1 and SNU-4139S3 sub-clones also presented analogous mutational profiles in parallel with drug responses. b. The similar pattern was observed in SNU-4631A series.

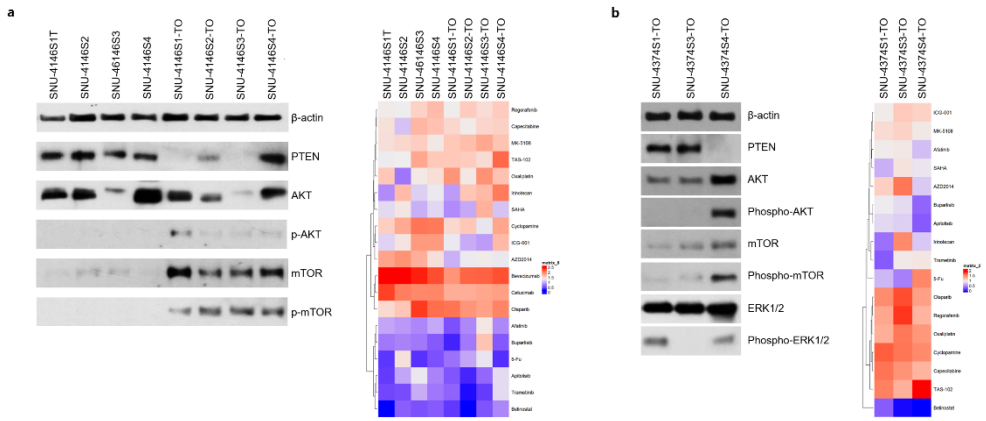


Figure 27. Expressional heterogeneity affects drug response. a. AKT-mTOR pathway was specifically activated in organoid samples. The response of AZD2014, the reagent inhibiting mTOR activation had good response to organoid groups. b. Similar pattern was observed in SNU-4374 series. AKT-mTOR pathway was explicitly up-regulated in SNU-4374S4-TO samples and AZD2014 showed better response.

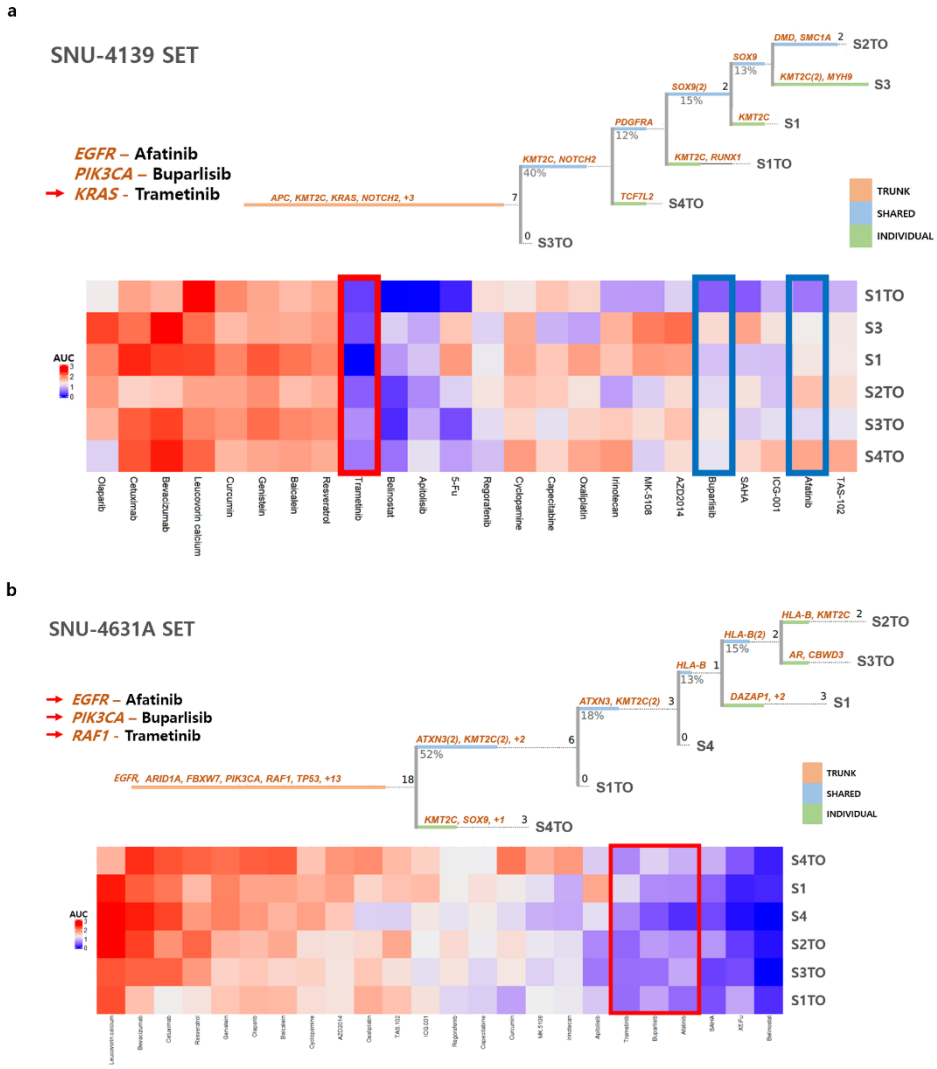


Figure 28. Targeting trunk mutation is effective to all sub-clones. All sub-clones of SNU-4139 series harbored KRAS mutation, and showed good response to trametinib (indicated with red box). In contrast, drugs targeting shared or individual mutations such as afatinib and buparlisib had varying response (indicated with blue box).

SNU-4631A series harbored EGFR, PIK3CA, and RAF1 mutations and showed good response to afatinib, buparlisib and trametinib (Figure 26b).

In sum, we have established CRC cell lines as well as organoids as a high-fidelity preclinical cancer model to deliver thorough understandings of tumor-related evolutionary trajectories in basic research and to be applied to personalized anti-cancer therapy in clinical. We have perceived the capability to reconstruct intra- and interpatient heterogeneity in CRC organoids and cell lines. Also, pathophysiological features of original CRC tumor were well recapitulated in CRC organoids. With these findings, we identified patient-derived colorectal organoids to be utilized to develop personalized anti-cancer therapy and prognostic biomarkers.

Acknowledgement

Soon-Chan Kim, Ja-Lok Ku, Young-Kyoung Shin, Ha-Young Seo, Ga-Hye Kim, Jae-Hyun Park, Ja Oh Lee, Myung Hyun Song, Seung-Yong Jeong, Ji-Won Park, Minjung Kim, Kyu Joo Park, Seung-Bum Ryoo, Jeong Mo Bae, and Bon-Kyoung Koo contributed to this study. This study is scheduled to be published afterward.

Reference

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2018;68(6):394-424.

2. Jung KW, Won YJ, Kong HJ, Lee ES. Cancer Statistics in Korea: Incidence, Mortality, Survival, and Prevalence in 2015. Cancer research and treatment : official journal of Korean Cancer Association. 2018;50(2):303-16.

3. Kolligs FT. Diagnostics and Epidemiology of Colorectal Cancer. Visceral medicine. 2016;32(3):158-64.

4. Xie YH, Chen YX, Fang JY. Comprehensive review of targeted therapy for colorectal cancer. Signal transduction and targeted therapy. 2020;5(1):22.

5. McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell. 2017;168(4):613-28.

6. Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. Nature genetics. 2014;46(3):225-33.

7. Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P, et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nat Med. 2015;21(7):751-9.

8. Murugaesu N, Wilson GA, Birkbak NJ, Watkins T, McGranahan N, Kumar S, et al. Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. Cancer Discov. 2015;5(8):821-31.

9. Zhang J, Fujimoto J, Zhang J, Wedge DC, Song X, Zhang J, et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. Science (New York, NY). 2014;346(6206):256-9.

10. de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science (New York, NY). 2014;346(6206):251-6.

11. Masoodi T, Siraj S, Siraj AK, Azam S, Qadri Z, Parvathareddy SK, et al. Genetic heterogeneity and evolutionary history of high-grade ovarian carcinoma and matched distant metastases. British journal of cancer. 2020;122(8):1219-30.

12. Schwarz RF, Ng CK, Cooke SL, Newman S, Temple J, Piskorz AM, et al. Spatial and temporal heterogeneity in high-grade serous ovarian cancer: a phylogenetic analysis. PLoS medicine. 2015;12(2):e1001789.

13. Haffner MC, Mosbruger T, Esopi DM, Fedor H, Heaphy CM, Walker DA, et al. Tracking the clonal origin of lethal prostate cancer. The Journal of clinical investigation. 2013;123(11):4918-22.

14. Gundem G, Van Loo P, Kremeyer B, Alexandrov LB, Tubio JMC, Papaemmanuil E, et al. The evolutionary history of lethal metastatic prostate cancer. Nature. 2015;520(7547):353-7.

15. Makohon-Moore AP, Zhang M, Reiter JG, Bozic I, Allen B, Kundu D, et al. Limited heterogeneity of known driver gene mutations among the metastases of individual patients with pancreatic cancer. Nature genetics. 2017;49(3):358-66.

16. Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. The New England journal of medicine. 2012;366(10):883-92.

17. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61(5):759-67.

18. JE IJ, Medema JP, Dekker E. Colorectal neoplasia pathways: state of the art. Gastrointestinal endoscopy clinics of North America. 2015;25(2):169-82.

19. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012;487(7407):330-7.

20. De Smedt L, Lemahieu J, Palmans S, Govaere O, Tousseyn T, Van Cutsem E, et al. Microsatellite instable vs stable colon carcinomas: analysis of tumour heterogeneity, inflammation and angiogenesis. British journal of cancer. 2015;113(3):500-9.

21. Perez-Villamil B, Romera-Lopez A, Hernandez-Prieto S, Lopez-Campos G, Calles A, Lopez-Asenjo JA, et al. Colon cancer molecular subtypes identified by expression profiling and associated to stroma, mucinous type and different clinical behavior. BMC cancer. 2012;12:260.

22. Schlicker A, Beran G, Chresta CM, McWalter G, Pritchard A, Weston S, et al. Subtypes of primary colorectal tumors correlate with response to targeted treatment in colorectal cell lines. BMC medical genomics. 2012;5:66.

23. Budinska E, Popovici V, Tejpar S, D'Ario G, Lapique N, Sikora KO, et al. Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. The Journal of pathology. 2013;231(1):63-76.

24. Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. PLoS medicine. 2013;10(5):e1001453.

25. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. Nat Med. 2015;21(11):1350-6.

26. Bijlsma MF, Sadanandam A, Tan P, Vermeulen L. Molecular subtypes in cancers of the gastrointestinal tract. Nature reviews Gastroenterology & hepatology. 2017;14(6):333-42.

27. Sveen A, Bruun J, Eide PW, Eilertsen IA, Ramirez L, Murumägi A, et al. Colorectal Cancer Consensus Molecular Subtypes Translated to Preclinical Models Uncover Potentially Targetable Cancer Cell Dependencies. Clinical cancer research : an official journal of the American Association for Cancer Research. 2018;24(4):794-806.

28. van de Wetering M, Francies HE, Francis JM, Bounova G, Iorio F, Pronk A, et al. Prospective derivation of a living organoid biobank of colorectal cancer patients. Cell. 2015;161(4):933-45.

29. Lau HCH, Kranenburg O, Xiao H, Yu J. Organoid models of gastrointestinal cancers in basic and translational research. Nature reviews Gastroenterology & hepatology. 2020;17(4):203-22.

30. Reiter JG, Makohon-Moore AP, Gerold JM, Bozic I, Chatterjee K, Iacobuzio-Donahue CA, et al. Reconstructing metastatic seeding patterns of human cancers. Nat Commun. 2017;8:14114.

31. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell. 2018;173(2):371-85.e18.

32. Sathirapongsasuti JF, Lee H, Horst BA, Brunner G, Cochran AJ, Binder S, et al. Exome sequencing-based copy-number variation and loss of heterozygosity detection: ExomeCNV. Bioinformatics (Oxford, England). 2011;27(19):2648-54.

33. Blokzijl F, Janssen R, van Boxtel R, Cuppen E. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. Genome medicine. 2018;10(1):33.

34. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature biotechnology. 2015;33(3):290-5.

35. Yao Y, Xu X, Yang L, Zhu J, Wan J, Shen L, et al. Patient-Derived Organoids Predict Chemoradiation Responses of Locally Advanced Rectal Cancer. Cell stem cell. 2020;26(1):17-26.e6.

36. Sachs N, de Ligt J, Kopper O, Gogola E, Bounova G, Weeber F, et al. A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. Cell. 2018;172(1-2):373-86.e10.

37. Cristobal A, van den Toorn HWP, van de Wetering M, Clevers H, Heck AJR, Mohammed S. Personalized Proteome Profiles of Healthy and

Tumor Human Colon Organoids Reveal Both Individual Diversity and Basic Features of Colorectal Cancer. Cell reports. 2017;18(1):263-74.

38. Dietlein F, Weghorn D, Taylor-Weiner A, Richters A, Reardon B, Liu D, et al. Identification of cancer driver genes based on nucleotide context. Nature genetics. 2020;52(2):208-18.

39. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature. 2014;505(7484):495-501.

40. Li SKH, Martin A. Mismatch Repair and Colon Cancer: Mechanisms and Therapies Explored. Trends in molecular medicine. 2016;22(4):274-89.

41. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nature genetics. 2013;45(10):1113-20.

42. Okugawa Y, Grady WM, Goel A. Epigenetic Alterations in Colorectal Cancer: Emerging Biomarkers. Gastroenterology. 2015;149(5):1204-25.e12.

43. Zhang X, Lan Y, Xu J, Quan F, Zhao E, Deng C, et al. CellMarker: a manually curated resource of cell markers in human and mouse. Nucleic acids research. 2019;47(D1):D721-d8.

44. Sewda K, Coppola D, Enkemann S, Yue B, Kim J, Lopez AS, et al. Cell-surface markers for colon adenoma and adenocarcinoma. Oncotarget. 2016;7(14):17773-89.

45. Beck TN, Korobeynikov VA, Kudinov AE, Georgopoulos R, Solanki NR, Andrews-Hoke M, et al. Anti-Müllerian Hormone Signaling Regulates Epithelial Plasticity and Chemoresistance in Lung Cancer. Cell reports. 2016;16(3):657-71.

46. Pandzic T, Rendo V, Lim J, Larsson C, Larsson J, Stoimenov I, et al. Somatic PRDM2 c.4467delA mutations in colorectal cancers control histone methylation and tumor growth. Oncotarget. 2017;8(58):98646-59.

47. Davidsen J, Larsen S, Coskun M, Gögenur I, Dahlgaard K, Bennett EP, et al. The VTI1A-TCF4 colon cancer fusion protein is a dominant negative regulator of Wnt signaling and is transcriptionally regulated by intestinal homeodomain factor CDX2. PloS one. 2018;13(7):e0200215.

48. Nakagawa S, Miyoshi N, Ishii H, Mimori K, Tanaka F, Sekimoto M, et al. Expression of CLDN1 in colorectal cancer: a novel marker for prognosis. Int J Oncol. 2011;39(4):791-6.

49. Cherradi S, Ayrolles-Torro A, Vezzo-Vié N, Gueguinou N, Denis V, Combes E, et al. Antibody targeting of claudin-1 as a potential

colorectal cancer therapy. Journal of experimental & clinical cancer research : CR. 2017;36(1):89.

50. Toiyama Y, Inoue Y, Yasuda H, Saigusa S, Yokoe T, Okugawa Y, et al. DPEP1, expressed in the early stages of colon carcinogenesis, affects cancer cell invasiveness. Journal of gastroenterology. 2011;46(2):153-63.

51. Eisenach PA, Soeth E, Röder C, Klöppel G, Tepel J, Kalthoff H, et al. Dipeptidase 1 (DPEP1) is a marker for the transition from low-grade to high-grade intraepithelial neoplasia and an adverse prognostic factor in colorectal cancer. British journal of cancer. 2013;109(3):694-703.

52. Murakami T, Kawada K, Iwamoto M, Akagami M, Hida K, Nakanishi Y, et al. The role of CXCR3 and CXCR4 in colorectal cancer metastasis. International journal of cancer. 2013;132(2):276-87.

53. Zhang L, Zhang L, Li H, Ge C, Zhao F, Tian H, et al. CXCL3 contributes to CD133(+) CSCs maintenance and forms a positive feedback regulation loop with CD133 in HCC via Erk1/2 phosphorylation. Scientific reports. 2016;6:27426.

54. Sadanandam A, Lyssiotis CA, Homicsko K, Collisson EA, Gibb WJ, Wullschleger S, et al. A colorectal cancer classification system that associates cellular phenotype and responses to therapy. Nat Med. 2013;19(5):619-25.

55. Karapetis CS, Khambata-Ford S, Jonker DJ, O'Callaghan CJ, Tu D, Tebbutt NC, et al. K-ras mutations and benefit from cetuximab in advanced colorectal cancer. The New England journal of medicine. 2008;359(17):1757-65.

56. Gamba S, Camaj P, Heinemann V, Laubender RP, Wang Y, Zhao Y, et al. Effect of KRAS exon 2 mutations on antitumor activity of afatinib and gefitinib. Anti-cancer drugs. 2015;26(4):371-8.

57. Tomić TT, Olausson J, Wilzén A, Sabel M, Truvé K, Sjögren H, et al. A new GTF2I-BRAF fusion mediating MAPK pathway activation in pilocytic astrocytoma. PloS one. 2017;12(4):e0175638.

58. Pauli C, Hopkins BD, Prandi D, Shaw R, Fedrizzi T, Sboner A, et al. Personalized In Vitro and In Vivo Cancer Models to Guide Precision Medicine. Cancer Discov. 2017;7(5):462-77.

59. Tuveson D, Clevers H. Cancer modeling meets human organoid technology. Science (New York, NY). 2019;364(6444):952-5.

60. Sato T, Stange DE, Ferrante M, Vries RG, Van Es JH, Van den Brink S, et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. Gastroenterology. 2011;141(5):1762-72.

61. Dijkstra KK, Cattaneo CM, Weeber F, Chalabi M, van de Haar J, Fanchi LF, et al. Generation of Tumor-Reactive T Cells by Co-culture of

Peripheral Blood Lymphocytes and Tumor Organoids. Cell. 2018;174(6):1586-98.e12.

62. Saito T, Niida A, Uchi R, Hirata H, Komatsu H, Sakimura S, et al. A temporal shift of the evolutionary principle shaping intratumor heterogeneity in colorectal cancer. Nat Commun. 2018;9(1):2884.

63. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. Nature. 2013;501(7467):338-45.

64. Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. Biochimica et biophysica acta. 2010;1805(1):105-17.

65. Thirlwell C, Will OC, Domingo E, Graham TA, McDonald SA, Oukrif D, et al. Clonality assessment and clonal ordering of individual neoplastic crypts shows polyclonality of colorectal adenomas. Gastroenterology. 2010;138(4):1441-54, 54.e1-7.

66. Walko CM, Lindley C. Capecitabine: a review. Clinical therapeutics. 2005;27(1):23-44.

67. Ciccolini J, Fina F, Bezulier K, Giacometti S, Roussel M, Evrard A, et al. Transmission of apoptosis in human colorectal tumor cells exposed to capecitabine, Xeloda, is mediated via Fas. Molecular cancer therapeutics. 2002;1(11):923-7.

68. Loo WT, Sasano H, Chow LW. Evaluation of therapeutic efficacy of capecitabine on human breast carcinoma tissues and cell lines in vitro. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie. 2007;61(9):553-7.

69. Wisniewska-Jarosinska M, Sliwinski T, Kasznicki J, Kaczmarczyk D, Krupa R, Bloch K, et al. Cytotoxicity and genotoxicity of capecitabine in head and neck cancer and normal cells. Molecular biology reports. 2011;38(6):3679-88.

70. De Roock W, Claes B, Bernasconi D, De Schutter J, Biesmans B, Fountzilas G, et al. Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. The Lancet Oncology. 2010;11(8):753-62.

71. Medico E, Russo M, Picco G, Cancelliere C, Valtorta E, Corti G, et al. The molecular landscape of colorectal cancer cell lines unveils clinically actionable kinase targets. Nat Commun. 2015;6:7002.

국문 초록

종양 내 이질성은 유전적으로 다양한 종양세포의 군집(Clonal heterogeneity)들이 섞여 있거나 혹은 공간적으로 분리되어 나타나게 되며, 이런 이질성으로 인하여 필연적으로 각각의 종양세포 군집 별로 상이한 형태의 세포 증식, 대사, 면역성, 전이능 등을 동반하게 된다. 따라서 종양 내 이질성은 유전적으로 다양한 세포 군집을 형성하며 이는 항암 치료에 살아남을 수 있는 소수 내성군집세포(Drug-tolerant cancer cells)를 제공할 수 있다. 따라서 종양 내 이질성을 이해하는 것은 항암치료의 표적 발굴을 위한 치료적 접근 뿐 아니라, 암환자의 예후 및 치료 반응성 예측과 같은 임상 진단 영역에서도 매우 중요하다. 현재 시행되고 있는 조직검사의 경우 단일 부위에서 채취된 샘플에 대해 분석이 이루어지며, 이러한 조직검사는 종양 내 이질성으로 인하여 부정확한 진단을 할 수 있다. 따라서 암의 재발 및 치료 내성을 진단하기 위해서는 다양한 부위 조직검사와 치료 전후에 검사 등, 다각도의 입체적인 검사를 필요로 하며 이와 관련된 새로운 검사 방법의 개발이 요구된다.

결장 직장암 (CRC)은 분자 및 임상 적 관점에서 매우 이질적인 질병이다. 최근 데이터에 따르면 결장 직장암은 공통 분자 하위 유형 (CMS1-4) 이라고 하는 4 개의 그룹으로 분리되며, 각각 고유 한 생물학 및 유전자 발현 패턴을 가지고 있다. 개선 된 아형 특이적 치료법을

개발하고 이들 아형의 분자 배선 및 기원에 대한 연구를 위해서는 이를 반영하는 생물학적 모델이 필요하다. 이 연구는 이질성을 반영하는 환자 유래 결장 직장암 세포주 및 오가노이드의 패널에서 CMS의 존재를 확인하도록 설계되었다. 이를 통한 종양 내 이질성을 이해한 치료적 접근은, 진단의 정확성 및 정밀성을 높이고 치료적 성공 가능성까지 높일 수 있는 기반이 될 수 있다.

주요어: 대장암, 오가노이드, 이질성, 항암제

학번: 2014-22004